



Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América

Dirección General de Estudios de Posgrado

Facultad de Ciencias Matemáticas

Unidad de Posgrado

**Un método de punto proximal escalarizado inexacto
para minimización multiobjetivo cuasi-convexa**

TESIS

Para optar el Grado Académico de Magíster en Matemática

Aplicada con mención en Matemática Computacional

AUTOR

Segundo CRUZADO ACUÑA

ASESOR

Erik Alex PAPA QUIROZ

Lima, Perú

2018



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Cruzado, S. (2018). *Un método de punto proximal escalarizado inexacto para minimización multiobjetivo cuasi-convexa*. [Tesis de maestría, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemáticas, Unidad de Posgrado]. Repositorio institucional Cybertesis UNMSM.

ACTA DE SUSTENTACIÓN DE TESIS DE GRADO ACADÉMICO DE MAGÍSTER

Siendo las 16:15 horas del día lunes 03 de diciembre del dos mil dieciocho, en el Auditorio de la Facultad de Ciencias Matemáticas, el Jurado Evaluador de Tesis, Presidido por el Dr. Renato Mario Benazic Tome e integrado por los siguientes miembros, Dra. María Natividad Zegarra Garay (Jurado Evaluador) Mg. Edinson Raúl Montoro Alegre (Jurado Informante); Mg. Jhelly Reynaluz Pérez Núñez (Jurado Evaluador), y el Dr. Erik Alex Papa Quiroz, como Miembro Asesor, se reunieron para la sustentación de la tesis titulada: «UN MÉTODO DE PUNTO PROXIMAL ESCALARIZADO INEXACTO PARA MINIMIZACIÓN MULTIOBJETIVO CUASI-CONVEXA» presentada por el Bachiller Segundo Cruzado Acuña para optar el Grado Académico de Magíster en Matemática Aplicada con mención en Matemática Computacional.

Luego de la exposición del graduando, los Miembros del Jurado hicieron las preguntas correspondientes, así como las observaciones e inquietudes acerca del trabajo de tesis, a las cuales el Bachiller Segundo Cruzado Acuña respondió con acierto y solvencia, demostrando pleno conocimiento del tema.

A continuación se realizó la calificación correspondiente, según tabla adjunta, resultando el Bachiller Segundo Cruzado Acuña aprobado con el calificativo de ...MUY BUENO... 13... (DIECIOCHO)

Habiendo sido aprobada la sustentación de la Tesis, el Jurado Evaluador recomienda para que el Consejo de Facultad apruebe el otorgamiento del Grado Académico de Magíster en Matemática Aplicada con mención en Matemática Computacional al Bachiller Segundo Cruzado Acuña.

Siendo las 17:45 horas, se levantó la sesión, firmando para constancia la presente Acta.



Dra. María Natividad Zegarra Garay
MIEMBRO



Dr. Renato Mario Benazic Tome
PRESIDENTE



Mg. Jhelly Reynaluz Pérez Núñez
MIEMBRO



Mg. Edinson Raúl Montoro Alegre
MIEMBRO



Dr. Erik Alex Papa Quiroz
MIEMBRO ASESOR

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS

ESCUELA DE POSGRADO

FACULTAD DE CIENCIAS MATEMÁTICAS

**UN MÉTODO DE PUNTO PROXIMAL ESCALARIZADO
INEXACTO PARA MINIMIZACIÓN MULTIOBJETIVO
CUASI-CONVEXA**

Tesis presentada a consideración del Cuerpo Docente de la de la Facultad de Ciencias Matemáticas de la Universidad Nacional Mayor de San Marcos, como parte de los requisitos para obtener el grado académico de Magíster en Matemática Aplicada, con Mención en Matemática Computacional.



Dr. Renato Mario Benazic Tome
Presidente



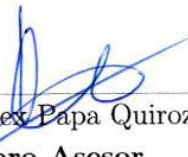
Dra. María Natividad Zegarra Garay
Miembro



Mg. Edinson Raúl Montoro Alegre
Miembro



Mg. Jhelly Reynaluz Pérez Núñez
Miembro



Dr. Erik Alex Papa Quiroz
Miembro Asesor

LIMA – PERÚ 2018

FICHA CATALOGRÁFICA

Un Método de Punto Proximal Escalarizado Inexacto para Minimización Multiobjetivo Cuasi-Conveza (Lima), 104p, (UNMSM, Magíster, Matemática Aplicada, 2018). Tesis, Universidad Nacional Mayor de San Marcos. Escuela de Posgrado, Facultad de Ciencias Matemáticas. Unidad de Posgrado.

1. Matemática Aplicada.

2. Algoritmo de Punto Proximal Escalarizado Inexacto.

I. UNMSM / EPG / FCM /UP.

II. Magíster (Serie).

DEDICATORIA

*Dedicado a mi familia, en
especial a la memoria de
Sergio Cruzado Cubas, mi
siempre querido, admirado
y recordado hermano.*

AGRADECIMIENTOS

Primeramente a mi Dios Todopoderoso por darme la vida y estar conmigo en todo momento.

A la memoria de mi hermano Sergio Cruzado Cubas, por impulsarme a través de su ejemplo a no dejar de estudiar.

Al Dr. Erik Papa Quiroz, por su colaboración en la elaboración y desarrollo de este trabajo, por apoyar y motivar las intenciones que tengo de seguir investigando, y por considerarme como miembro del grupo de investigación que preside. Mis respeto y admiración para usted.

A mis compañeros de clase, por acompañarme en el proceso, y alentarme en el desarrollo de la presente tesis.

¡Muchas gracias a todos!

Índice general

Resumen	vii
Abstract	xi
1. INTRODUCCIÓN	1
1.1. Situación problemática	1
1.2. Formulación del problema	3
1.2.0.1. Problema general	3
1.2.0.2. Problemas específicos	3
1.3. Justificación de la investigación	3
1.3.1. Objetivos	4
1.3.1.1. Objetivo general	4
1.3.1.2. Objetivos específicos	4
2. MARCO TEÓRICO	5
2.1. Elementos básicos del análisis real	5
2.1.1. Producto interno en \mathbb{R}^n	5
2.1.2. La norma en \mathbb{R}^n	6
2.1.3. Bolas, sucesiones y series en Espacios Euclidianos	7
2.1.3.1. Bolas en \mathbb{R}^n	7
2.1.3.2. Sucesiones en \mathbb{R}^n	7
2.1.3.3. Subsucesiones	9
2.1.3.4. Convergencia de sucesiones	9
2.1.3.5. Series infinitas	10
2.1.4. Conjuntos abiertos y cerrados	11
2.1.4.1. Interior de un conjunto	11
2.1.4.2. Clausura de un conjunto	12
2.1.4.3. Puntos de acumulación	13
2.1.5. Funciones continuas	14
2.1.6. Funciones de Lipschitz	15
2.2. Elementos básicos del análisis convexo	15
2.2.1. Generalidades de las funciones de \mathbb{R}^n en $\mathbb{R} \cup \{+\infty\}$	15
2.2.2. Convexidad y cuasi-convexidad de funciones	18
2.2.3. Derivadas Direccionales, Subgradiientes y Subdiferenciales	20
2.3. Compacidad y existencia de minimizadores	23
2.3.1. Óptimos locales y globales	23
2.3.2. Resultados de existencia	24

3. SUBDIFERENCIALES GENERALIZADOS Y DIRECCIONES DE DESCENSO	27
3.1. Subdiferenciales Generalizados	27
3.1.1. Subdiferencial de Fréchet y en el límite	28
3.1.2. Subdiferencial de Clarke	30
3.2. Direcciones de descenso	39
4. TEORÍA DE OPTIMIZACIÓN MULTIOBJETIVO	41
4.1. Ordenes de preferencia	43
4.2. Conos y ordenes relacionados	47
4.3. Eficiencia, débil eficiencia y no dominancia de soluciones	52
5. MÉTODO DE PUNTO PROXIMAL ESCALARIZADO INEXACTO (MPPEI)	63
5.1. Antecedentes de la investigación	63
5.2. El algoritmo	65
5.2.1. Existencia de los iterados	67
5.2.2. Resultados de convergencia	69
5.3. Variantes inexactas del algoritmo MPPEI para minimización multiobjetivo	72
5.3.1. Algoritmo MPPEI1 y resultados de convergencia	72
5.3.2. Algoritmo MPPEI2 y convergencia de resultados	77
6. TASA DE CONVERGENCIA DEL ALGORITMO MPPEI2	81
6.1. Análisis de la tasa de convergencia	82
6.2. Experimentación numérica	85
7. METODOLOGÍA	95
7.1. Tipo y diseño de investigación	95
7.2. Unidad de análisis	95
7.3. Población de estudio	95
7.4. Técnicas de recolección de datos	95
7.5. Análisis e interpretación de la información	96
8. RESULTADOS Y DISCUSIÓN	97
9. CONCLUSIONES Y FUTURAS INVESTIGACIONES	99
9.1. Conclusiones	99
9.2. Futuras investigaciones	100
REFERENCIAS BIBLIOGRÁFICAS	100

RESUMEN

Autor:	Segundo Cruzado Acuña
Asesor:	Dr. Erik Papa Quiroz
Grado obtenido:	Magíster en matemática aplicada
Mención:	Matemática Computacional

En la presente tesis presentamos un método de punto proximal escalarizado inexacto para resolver problemas irrestrictos de minimización multiobjetivo cuasi-convexos definidos en espacios Euclidianos, donde las funciones vectoriales son localmente Lipschitz.

Bajo algunas hipótesis naturales, se prueba que la sucesión generada por el método está bien definida y converge globalmente. Seguidamente proporcionando al método propuesto dos criterios de error, se obtienen dos variantes del mismo, y se prueba que las sucesiones generadas por cada una de esas variantes, convergen hacia un punto crítico Pareto-Clarke del problema; también se prueba que al dotar a la función vectorial de ciertas condiciones, la tasa de convergencia de uno de éstos métodos es lineal y superlineal. Finalmente para validar el método propuesto y los resultados encontrados, se presentan algunos experimentos computacionales.

Palabras clave:

- Método de Punto proximal
- Función Cuasi-convexa
- Optimización Multiobjetivo
- Subdiferencial de Clarke
- Subdiferencial de Fréchet
- Tasa de Convergencia
- Convergencia Lineal y Superlineal

ABSTRACT

Author:	Segundo Cruzado Acuña
Advisor:	Dr. Erik Papa Quiroz
Obtained Degree:	Magister in applied mathematics
Emphasis:	Computational Mathematics

In this paper, we present an inexact scalarized proximal point method to solve unconstrained quasiconvex multiobjective minimization problems defined in Euclidean spaces, where the vector functions are locally Lipschitz

Under some natural assumptions, it is proved that the sequence generated by the method is well defined and converges globally. Next, introducing two error criteria on the method, two variants are obtained, and it is proved that the sequences generated by each one of those variants, converge to a Pareto-Clarke critical point of the problem. Then providing these methods with two error criteria, two versions of it are obtained, and it is proved that any accumulation point of the sequences that they generate, converge toward a Pareto-Clarke critical point of the problem; it is also proved that by providing the vector function with certain conditions, the convergence rate of one of these methods is linear and superlinear. Finally, to validate the proposed method and the results found, some computational experiments are presented.

Keywords:

- Proximal Point Method
- Quasiconvex Function
- Multiobjective Optimization
- Clarke Subdifferential
- Fréchet Subdifferential
- Convergence Rate
- Linear and Superlinear Convergence

Índice de tablas

4.1.	Características de los cuatro automóviles (caso discreto).	41
4.2.	Algunos órdenes en \mathbb{R}^p	47
4.3.	Fuente: Ehrgott (2005), Tabla 1.2, página12.	47
4.4.	Características de los cuatro automóviles de la variante bi-objetivo del Ejemplo 4.0.1 (caso discreto).	55
4.7.	Vértices del conjunto factible X en el espacio variable, e imágenes de los mismos en el espacio objetivo.	59

Índice de figuras

2.1. (a) Epígrafo y (b) conjunto de nivel inferior de una función no convexa.	16
2.2. La función f es sci, sin embargo no es continua.	17
2.3. d_1 y d_2 son los subgradientes de f en x_0	21
2.4. La función valor absoluto.	21
2.5. Óptimos locales y globales.	24
2.6. La función exponencial tiene ínfimo pero no mínimo.	25
3.1. Gráfica de la función $f(x) = x^2 \text{sen}(1/x)$	31
4.1. Gráfica de las funciones objetivo del Ejemplo 4.0.2.	43
4.2. Ilustración de dos conos.	48
4.3. Ilustración del orden de Pareto “ \leq ” en \mathbb{R}^2 . Fuente Tablada (2011) - Figura 2.4.	51
4.4. Ilustración de las definiciones de soluciones eficientes.	54
4.5. Espacio objetivo del Ejemplo 4.3.1.	55
4.6. (a) Espacio objetivo y (b) puntos no dominados del Ejemplo 4.3.2.	57
4.7. Regiones de dominancia del punto $P_1(0, 4)$	58
4.8. (a) Conjunto factible X en el espacio variable. (b) Conjunto factible Y y conjunto no dominado Y_N en el espacio objetivo.	59
4.9. (a) Y_N es vacío, pero Y_{wN} no lo es. (b) Punto no dominado, y puntos débilmente no dominados.	61
4.10. Regiones de dominancia de los conjuntos Y_N y Y_{wN}	62
6.1. Gráfica de las funciones objetivo del Ejemplo 6.2.3.	92

NOTACIONES Y SÍMBOLOS

Notación	Explicación
X	conjunto factible en el espacio de decisión $X \subseteq \mathbb{R}^n$.
$Y = F(X)$	conjunto factible en el espacio objetivo $Y \subset \mathbb{R}^p$.
C	cono.
$x = (x_1, \dots, x_n)$	vector de variables de decisión.
$y = (y_1, \dots, y_p)$	vector de valores de la función objetivo.
$F = (F_1, \dots, F_p)$	vector de funciones objetivo, donde $F : \mathbb{R}^n \rightarrow \mathbb{R}^p$.
$y^1 < y^2$	$y_k^1 < y_k^2$, para $k = 1, \dots, p$.
$y^1 \leq y^2$	$y_k^1 \leq y_k^2$, para $k = 1, \dots, p$.
$y^1 \preceq y^2$	$y_k^1 \leq y_k^2$, pero $y^1 \neq y^2$.
\mathbb{R}^n	$\{x = (x_1, \dots, x_n) \in \mathbb{R}^n : x_i \in \mathbb{R}, \forall i = 1, \dots, p\}$ denota el espacio Euclidiano n-dimensional. Cuando se trabaja con matrices x , es representado como un vector columna.
$\overline{\mathbb{R}}$	$\mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$ conjunto de los números reales extendidos.
$\mathbb{R}_{>}^p = \mathbb{R}_{++}^p$	$\{y \in \mathbb{R}^p : y > 0\}$ conjunto de los números reales positivos.
$\mathbb{R}_{\geq}^p = \mathbb{R}_{\geq}^p \setminus \{0\}$	$\{y \in \mathbb{R}^p : y \geq 0\}$.
$\mathbb{R}_{\geq}^p = \mathbb{R}_+^p$	$\{y \in \mathbb{R}^p : y \geq 0\}$ (ortante no negativo) conjunto de los números reales no negativos.
$x \succ y$	x es preferido a y .
$x \succ y$	x es estrictamente preferido a y .
$int(X)$	interior del conjunto X .
$ri(X)$	interior relativo del conjunto X .
$\overline{X} = cl(X)$	cerradura o clausura del conjunto X .
$Fr(Y)$	frontera del conjunto Y .
$X^c = \mathbb{R}^n \setminus X$	complemento del conjunto X .
$x \in X \setminus Y$	significa que $x \in X$ y $x \notin Y$.

$x^k \rightarrow x$	denota que $\lim_{k \rightarrow +\infty} x^k = x$.
$t \downarrow 0$	denota que $t \rightarrow 0^+$.
$t \searrow 0$	denota que t converge a cero de forma decreciente.
$t \nearrow 0$	denota que t converge a cero de forma creciente.
∇f	gradiente de f .
$f'(x, d)$	derivada direccional (en el sentido usual) de f en x en la dirección $d \in \mathbb{R}^n$.
$\partial f(x)$	subdiferencial de f en x (en el sentido del análisis convexo) .
$\widehat{\partial} f(x)$	conjunto de subgradientes regulares (también llamado subdiferencial de Fréchet) de f en x .
$\partial^{Lim} f(x)$	conjunto de subgradientes generalizados (también llamado subdiferencial en el límite) de f en x .
$f^0(x, d)$	derivada direccional de Clarke de f en x en la dirección del vector d .
$\partial^0 f(x)$	subdiferencial de Clarke de f en x .
$\langle x, y \rangle = x^T y$	$\sum_{i=1}^n x_i y_i$ producto interior de $x, y \in \mathbb{R}^n$.
$\ x\ = \ x\ _2$	$\sqrt{\langle x, x \rangle}$ norma Euclidiana.
$\ x\ _1 = \sum_{i=1}^n x_i $	norma 1.
$B(\bar{x}, \epsilon)$	$\{x \in \mathbb{R}^n : \ x - \bar{x}\ < \epsilon\}$ bola abierta con centro el punto \bar{x} y radio ϵ .
$\delta_C(\cdot)$	función indicatriz del conjunto no vacío $C \subset \mathbb{R}^n$.
$N_C(\bar{x})$	$\{d \in \mathbb{R}^n : \langle d, x - \bar{x} \rangle \leq 0, \forall x \in C\}$ cono normal al punto \bar{x} relacionado al conjunto $C \subset \mathbb{R}^n$.
$\text{epi}(f)$	$\{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R} : f(x) \leq \alpha\}$ epígrafo de la función f .
$L_f(\alpha)$	$\{x \in \mathbb{R}^n : f(x) \leq \alpha\}$ conjunto de nivel inferior de f .
$\liminf_{k \rightarrow +\infty} f(x^k)$	$\sup_{n \geq 0} \inf_{k \geq n} \{f(x^k)\}$ limite inferior de la sucesión.
$\limsup_{k \rightarrow +\infty} f(x^k)$	$\inf_{n \geq 0} \sup_{k \geq n} \{f(x^k)\}$ limite superior de la sucesión.
$\limsup_{\substack{y \rightarrow x \\ t \downarrow 0}} f(y)$	$\inf_{\delta > 0} \sup_{y \in B(x, \delta), t \in (0, \delta)} f(y) = \inf_{\delta > 0} \sup_{\ y-x\ +t < \delta} f(y)$ limite superior de una función localmente Lipschitz.
$z \in \underset{y \in X}{\operatorname{argmin}} \{g(y)\}$	implica que $g(z) \leq g(y)$ para todo $y \in X$, donde $g : \mathbb{R}^n \supset X \rightarrow \mathbb{R}$.

Capítulo 1

INTRODUCCIÓN

1.1. Situación problemática

En el campo de la matemática aplicada existen diversas áreas que se encargan de dar solución a problemas vinculados a las ciencias e ingenierías, una de ellas es la optimización matemática, la cuál nos brinda los criterios y métodos necesarios para encontrar la mejor solución a un problema de entre un conjunto de posibles alternativas. El modelo general de optimización está denotado por:

$$\underset{s.a \ x \in X}{opt} \{F(x) : x \in X\} \equiv \underset{s.a \ x \in X}{opt} F(x), \quad (1.1)$$

donde *opt.* significa maximizar o minimizar la función objetivo $F : \mathbb{R}^n \rightarrow \mathbb{R}^p$, $p \in \mathbb{Z}^+$, y $X \subseteq \mathbb{R}^n$ es el conjunto de restricciones del problema. Dependiendo de la naturaleza de F , y de la forma como esté definido el conjunto X , tendremos por ejemplo problemas de optimización mono-objetivo (donde $F = f : \mathbb{R}^n \rightarrow \mathbb{R}$), multiobjetivo ($F : \mathbb{R}^n \rightarrow \mathbb{R}^p$, $p \geq 2$), restrictos (si X está definido por restricciones de caja, o por restricciones de desigualdad y/o igualdad), irrestrictos (si $X = \mathbb{R}^n$), cuasi-convexos¹, convexos, lineales, etc. Así con la combinación de las clasificaciones anteriores, por ejemplo podemos tener un problema irrestricto de minimización multiobjetivo cuasi-convexo, problema que será abordado para los fines de la presente tesis.

La complejidad de algunos problemas de optimización, a motivado a muchos investigadores a desarrollar métodos que permitan resolver el problema propuesto. Uno de éstos es el Método del Punto Proximal (MPP), introducido por Martinet (1970), el cuál es un método iterativo, que se caracteriza por resolver en cada iteración un problema más simple de optimización, de tal forma que paso a paso las iteraciones se aproximan a la solución buscada.

Dicho método surgió inicialmente para resolver problemas de minimización convexos, y posteriormente fue utilizado por Rockafellar (1976) para encontrar ceros de operadores monótonos maximales. Fue tal la acogida que tuvieron estos algoritmos, que posteriormente fueron extendidos hacia las desigualdades variacionales monótonas y no monótonas, y en particular hacia los problemas no convexos de

¹Un problema de optimización es cuasi-convexo, si tanto las funciones objetivo, como las funciones que definen las restricciones (en el caso restricto), son cuasi-convexas. Decimos que $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es cuasi-convexa, si $f(\lambda x + (1 - \lambda)y) \leq \max\{f(x), f(y)\}$, $\forall \lambda \in [0, 1]$, $\forall x, y \in \mathbb{R}^n$.

minimización, lo cual no solo se debe a la simple extensión de la teoría de convergencia del método, sino también a las múltiples aplicaciones en diversas áreas de las ciencias e ingenierías, aplicaciones que han permitieron que en las últimas décadas dichos algoritmos también sean extendidos hacia los problemas de minimización cuasi-convexos.

La necesidad de construir algoritmos de punto proximal computacionalmente más implementables, a hecho que éstos sean extendidos hacia sus versiones inexactas, lo cual se justifica en la dificultad que por lo general se tiene para encontrar un valor x^k que resuelva de forma exacta los subproblemas generados por el Método del Punto Proximal. En ese sentido se consideran errores de aproximación e^k , los cuales al cumplir determinadas propiedades, permiten generalizar las propiedades de convergencia de los algoritmos de búsqueda exacta, ver por ejemplo Papa Quiroz y Oliveira (2012); Papa Quiroz et al. (2015).

En el año 2016 Apolinario, Papa Quiroz y Oliveira, trabajando con funciones multiobjetivo, y asumiendo que éstas son cuasi-convexas, localmente Lipschitz, y están acotadas inferiormente por el cero, presentan el artículo **“A scalarization proximal point method for quasiconvex multiobjective minimization”** Apolinário et al. (2016), el cual tiene como objeto resolver el problema de minimización (1.1) en su versión multiobjetivo. La iteración que consideran es la siguiente. Dado un punto inicial arbitrario $x^k \in \mathbb{R}^n$, encontrar un $x^{k+1} \in \Omega_k$, tal que:

$$0 \in \partial^0 \left(\langle F(\cdot), z_k \rangle + \frac{\lambda_k}{2} \|\cdot - x^k\|^2 \right) (x^{k+1}) + \mathcal{N}_{\Omega_k}(x^{k+1}), \quad (1.2)$$

donde ∂^0 es el subdiferencial de Clarke (ver Sección 3.1), $\Omega_k = \{x \in \mathbb{R}^n : F(x) \preceq F(x^k)\}$, $\lambda_k > 0$, $\{z_k\} \subset \mathbb{R}_{\geq}^p \setminus \{0\}$, $\|z_k\| = 1$, y $\mathcal{N}_{\Omega_k}(x^{k+1})$ es el cono normal a Ω_k en x^{k+1} . Los autores prueban que la sucesión $\{x^{k+1}\}$ generada por el método está bien definida, y además obtienen la convergencia global a un punto crítico Pareto-Clarke, y cuando F es convexa, obtienen la convergencia a una solución débil de Pareto.

La versión inexacta del método anterior en el espacio de las variedades de Hadamard (variedades Riemannianas con curvatura seccional no positiva), actualmente está siendo estudiada, sin embargo aún no existen publicaciones al respecto, ver por ejemplo la tesis doctoral **“Método do ponto proximal inexato para minimização quase-convexa em variedades de Hadamard”**, presentada por Baygorrea (2017), en la Universidad Federal de Rio de Janeiro. Entonces, al no existir en los espacios Euclidianos trabajos teóricos relacionados con la extensión de dicho método hacia su versión inexacta, nos proponemos realizar dicho trabajo.

1.2. Formulación del problema

1.2.0.1. Problema general

¿Cómo extender a su versión inexacta el Método de Punto Proximal Escalarizado (MPPE) para minimización multiobjetivo cuasi-convexa?.

1.2.0.2. Problemas específicos

1. ¿Qué condiciones debe tener el error de aproximación para que el MPPE pueda ser extendido a su versión inexacta?.
2. ¿Cuáles serán las condiciones de convergencia para el MPPE inexacto?.
3. ¿De qué manera se medirá la velocidad de convergencia del MPPE inexacto?.
4. ¿Cómo resolver un caso particular con el MPPE inexacto?.

1.3. Justificación de la investigación

El desarrollo del presente proyecto se justifica por lo siguiente.

1. Por la existencia de investigaciones que dan fe que las iteraciones inexactas (en comparación con los de búsqueda exacta) son muy prometedoras y efectivas para encontrar soluciones aproximadas de problemas de optimización. Efectividad traducida en menos cantidad de iteraciones, rápida convergencia, y por ende menos costo computacional. Entre los reportes cabe resaltar los artículos Papa Quiroz y Oliveira (2012); Papa Quiroz et al. (2015), en el último de los cuáles los autores (quienes trabajando en ambos casos con funciones cuasi-convexas semi-continuas inferiores), presentan una versión inexacta del primero de ellos. Otros trabajos para funciones cuasi-convexas diferenciables pero que hacen uso de iteraciones exactas también fueron desarrollados por Pan y Chen (2007), Chen y Pan (2008), Cunha et al. (2010) y Souza et al. (2010).
2. Por la no existencia de publicaciones relacionados con la convergencia global de la versión inexacta del MPPE para minimización de funciones multiobjetivo cuasi-convexas definidas en espacios Euclidianos. En éste punto cabe recalcar que dicho algoritmo en el espacio de las variedades de Hadamard, está siendo estudiado Baygorrea (2017).
3. Por el gran interés que en las últimas décadas ha despertado la extensión de los algoritmos de punto proximal (APP) de su versión exacta, a una inexacta, de estar restringidos, a irrestrictos, de ser convexos a no convexos, lineales, a no lineales, etc. Dicho interés se debe no solo a la simple extensión de la teoría de convergencia global del método, sino también a las múltiples aplicaciones en diversas áreas de las ciencias e ingeniería, ver por ejemplo los trabajos de Kaplan y Tichatschke (1998), Pennanen (2002), Chen y Pan

(2008), Attouch y Bolte (2009) y Attouch et al. (2013).

En particular por la reciente atención que vienen recibiendo los problemas de minimización cuasi-convexos, debido a su amplia gama de aplicaciones. Por ejemplo en teoría de localización dos Santos Gromicho (2013), en programación fraccional, y especialmente en teoría económica, ver por ejemplo Takayama (1995) y Mas-Colell et al. (1995). Algunos artículos relacionados a ellos son los siguientes: Goudou y Munier (2009), Souza et al. (2010), quiroz2012proximal, entre otros.

4. Finalmente por que esta investigación va a permitir que a futuro pueda desarrollar algoritmos de optimización más eficientes, los cuáles fueron la motivación para el desarrollo del presente trabajo.

1.3.1. Objetivos

1.3.1.1. Objetivo general

Extender a su versión inexacta el MPPE para minimización multiobjetivo cuasi-convexa.

1.3.1.2. Objetivos específicos

1. Determinar las condiciones que el error de aproximación debe tener para que el MPPE pueda ser extendido a su versión inexacta.
2. Extender las propiedades de convergencia del MPPE, para minimización de funciones cuasi-convexas.
3. Determinar la velocidad de convergencia del MPPE inexacto.
4. Aplicar el MPPE inexacto a un caso particular.

Capítulo 2

MARCO TEÓRICO

El desarrollo de éste trabajo está basado en teoría relacionada a los métodos punto proximal para minimización multiobjetivo cuasi-convexa, y en los procedimientos que permiten la extensión de los mismos hacia su versiones inexacta.

2.1. Elementos básicos del análisis real

2.1.1. Producto interno en \mathbb{R}^n

Definición 2.1.1. Sean $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n) \in \mathbb{R}^n$, el producto interno de x e y denotado por $\langle x, y \rangle$ es el número real

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i.$$

Las principales propiedades del producto interno son resumidas en el siguiente teorema.

Teorema 2.1.1. Se cumplen las siguientes propiedades:

1. $\langle x, x \rangle \geq 0$, para todo $x \in \mathbb{R}^n$.
2. $\langle x, x \rangle = 0$, si y sólo si $x = 0$.
3. $\langle x, y \rangle = \langle y, x \rangle$, para todo $x, y \in \mathbb{R}^n$.
4. $\langle x + x', y \rangle = \langle x, y \rangle + \langle x', y \rangle$, para todo $x, x', y \in \mathbb{R}^n$.
5. $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$, para todo $\alpha \in \mathbb{R}$ y para todo $x, y \in \mathbb{R}^n$.

Demostración. Véase Benazic (2000), página 5. □

2.1.2. La norma en \mathbb{R}^n

Definición 2.1.2. Sea $x \in \mathbb{R}^n$, la norma de x , denotada por $\|x\|$, es definida como

$$\|x\| = \sqrt{\langle x, x \rangle}$$

Observación 2.1.1. La norma del vector x se interpreta geométricamente como la distancia del punto x al punto θ (origen de coordenadas). Esta norma es llamada **norma Euclídiana**.

Teorema 2.1.2. (Desigualdad de Cauchy-Schwarz)

1. $|\langle x, y \rangle| \leq \|x\| \|y\|$, para todo $x, y \in \mathbb{R}^n$.
2. $|\langle x, y \rangle| = \|x\| \|y\|$, si y sólo si existe $\alpha \in \mathbb{R}$ tal que $x = \alpha y$.

Demostración. Véase Benazic (2000), página 8. □

Teorema 2.1.3. Se cumplen las siguientes propiedades:

1. $\|x\| \geq 0$, para todo $x \in \mathbb{R}^n$.
2. $\|x\| = 0$, si y sólo si $x = 0$.
3. $\|x + y\| \leq \|x\| + \|y\|$, para todo $x, y \in \mathbb{R}^n$ (Desigualdad Triangular).
4. $\|\alpha x\| = |\alpha| \|x\|$, para todo $x \in \mathbb{R}^n$ y para todo $\alpha \in \mathbb{R}$.

Demostración. Véase Benazic (2000), página 9. □

Proposición 2.1.1. Se cumplen las siguientes propiedades:

1. $|||x| - |y|| \leq \|x - y\|$, para todo $x, y \in \mathbb{R}^n$.
2. $\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$, para todo $x, y \in \mathbb{R}^n$ (Identidad del Paralelogramo).
3. $\|x + y\|^2 - \|x - y\|^2 = 4\langle x, y \rangle$, para todo $x, y \in \mathbb{R}^n$ (Identidad de Polarización).

Demostración. Véase Benazic (2000), página 10. □

2.1.3. Bolas, sucesiones y series en Espacios Euclidianos

2.1.3.1. Bolas en \mathbb{R}^n

La norma Euclidiana en \mathbb{R}^n estudiada en la sección anterior, nos permite definir algunas nociones geométricas fundamentales.

Definición 2.1.3. Sean $x \in \mathbb{R}^n$ y $r > 0$.

1. La bola abierta de centro \bar{x} y radio r es el conjunto definido por:

$$B(\bar{x}, r) = B_r(\bar{x}) = \{x \in \mathbb{R}^n : \|x - \bar{x}\| < r\}.$$

2. La bola cerrada de centro \bar{x} y radio r es el conjunto definido por:

$$B[\bar{x}, r] = B_r[\bar{x}] = \{x \in \mathbb{R}^n : \|x - \bar{x}\| \leq r\}.$$

2.1.3.2. Sucesiones en \mathbb{R}^n

Definición 2.1.4. Una sucesión en \mathbb{R}^n es una función $x : \mathbb{N} \rightarrow \mathbb{R}^n$ que a cada número natural k le asocia el vector $x(k) = x_k \in \mathbb{R}^n$ llamado k -ésimo término de la sucesión.

Las sucesiones serán denotadas por $\{x_1, x_2, \dots, x_k, \dots\}$, $\{x_k\}_{k \in \mathbb{N}}$, o simplemente por $\{x_k\} \subseteq \mathbb{R}^n$.

Definición 2.1.5. Sea $\{x_k\} \subseteq \mathbb{R}^n$, decimos que:

1. $\{x_k\}$ es creciente, lo cual denotamos por $x_k \nearrow$, si $x_k \leq x_{k+1} \quad \forall k \geq 1$.
2. $\{x_k\}$ es decreciente, lo cual denotamos por $x_k \searrow$, si $x_k \geq x_{k+1} \quad \forall k \geq 1$.

Una sucesión es llamada monótona cuando es creciente o decreciente.

Definición 2.1.6. Sea $\{x_k\} \subseteq \mathbb{R}^n$, decimos que $x \in \mathbb{R}^n$ es el límite de la sucesión $\{x_k\}$ cuando k tiende al infinito, lo que denotamos por $\lim_{k \rightarrow +\infty} x_k = x$ si y sólo si para todo $\epsilon > 0$, existe $k_0 \in \mathbb{N}$ tal que si $k \geq k_0$, entonces $\|x_k - x\| < \epsilon$. Simbólicamente:

$$\lim_{k \rightarrow +\infty} x_k = x \Leftrightarrow \forall \epsilon > 0, \exists k_0 \in \mathbb{N}; n \geq n_0 \implies \|x_k - x\| < \epsilon.$$

Definición 2.1.7. Sea $\{x_k\} \subseteq \mathbb{R}^n$, decimos que:

1. $\{x_k\}$ es una **sucesión convergente**, si y sólo si existe $x \in \mathbb{R}^n$ tal que $\lim_{k \rightarrow +\infty} x_k = x$.
2. $\{x_k\}$ es una sucesión divergente si y sólo si no es convergente.

Ejemplo 2.1.1. Para determinar la convergencia o divergencia de una sucesión, por ejemplo de $\{\frac{k+1}{2k+1}\}_{k \geq 1}$, basta calcular el límite de la misma, es decir:

$$\lim_{k \rightarrow +\infty} x_k = \lim_{k \rightarrow +\infty} \frac{k+1}{2k+1} = \lim_{k \rightarrow +\infty} \frac{1 + \frac{1}{k}}{2 + \frac{1}{k}} = \lim_{k \rightarrow +\infty} \frac{1+0}{2+0} = \frac{1}{2}. \quad (2.1)$$

Por lo tanto $\{\frac{k+1}{2k+1}\}$ es convergente. \square

Proposición 2.1.2. (Unicidad del límite) El límite de una sucesión convergente de vectores es único. Es decir si $\lim_{k \rightarrow +\infty} x_k = a$ y $\lim_{k \rightarrow +\infty} x_k = b$, entonces $a = b$.

Demostración. Véase Benazic (2000), página 20. \square

Corolario 2.1.1. Si $\{x_k\}, \{y_k\} \subseteq \mathbb{R}^n$ y $\alpha_k \subseteq \mathbb{R}$ son sucesiones convergentes, entonces $\{x_k \pm y_k\}, \{\alpha_k x_k\} \subseteq \mathbb{R}^n$, $\{\langle x_k, y_k \rangle\}, \{\|x_k\|\}, \{d(x_k, y_k)\} \subseteq \mathbb{R}$ también son sucesiones convergentes, y además se cumple:

1. $\lim_{k \rightarrow +\infty} (x_k \pm y_k) = \lim_{k \rightarrow +\infty} (x_k) \pm \lim_{k \rightarrow +\infty} (y_k).$
2. $\lim_{k \rightarrow +\infty} (\alpha_k x_k) = \left(\lim_{k \rightarrow +\infty} \alpha_k \right) \left(\lim_{k \rightarrow +\infty} x_k \right).$
3. $\lim_{k \rightarrow +\infty} \langle x_k, y_k \rangle = \langle \lim_{k \rightarrow +\infty} x_k, \lim_{k \rightarrow +\infty} y_k \rangle.$
4. $\lim_{k \rightarrow +\infty} \|x_k\| = \left\| \lim_{k \rightarrow +\infty} x_k \right\|.$
5. $\lim_{k \rightarrow +\infty} d(x_k, y_k) = d \left(\lim_{k \rightarrow +\infty} x_k, \lim_{k \rightarrow +\infty} y_k \right).$

Demostración. Véase Benazic (2000), página 21. \square

Definición 2.1.8. Sea $\{x_k\} \subseteq \mathbb{R}^n$, decimos que $\{x_k\}$ es una **sucesión acotada** si existe una constante positiva C tal que $\|x_k\| \leq C$ para todo $k \in \mathbb{N}$.

Proposición 2.1.3. Si $\{x_k\} \subseteq \mathbb{R}^n$ es una sucesión convergente, entonces $\{x_k\}$ es acotada.

Demostración. Véase Benazic (2000), página 21. \square

2.1.3.3. Subsucesiones

Definición 2.1.9. Sea $\{x_k\} \subseteq \mathbb{R}^n$ y $j : \mathbb{N} \rightarrow \mathbb{N}$ una función creciente (es decir $j_1 < j_2 < \dots < j_k < \dots$). La composición $x \circ j : \mathbb{N} \rightarrow \mathbb{R}^n$ que a cada número natural k le asocia el vector $x \circ j(k) = x(j(k)) = x_{j_k}$ es llamada subsucesión de $\{x_k\}$.

La notación $\{x_{j_k}\} \subseteq \{x_k\}$ significa que $\{x_{j_k}\}$ es una subsucesión de $\{x_k\}$.

Proposición 2.1.4. Sea $\{x_k\} \subseteq \mathbb{R}^n$. Si $\lim_{k \rightarrow +\infty} x_k = x$ entonces toda subsucesión de $\{x_k\}$ converge hacia x .

Demostración. Véase Benazic (2000), página 22. □

Teorema 2.1.4. (Bolzano-Weierstrass). Toda sucesión acotada en \mathbb{R}^n posee una subsucesión convergente.

Demostración. Véase Benazic (2000), página 22. □

Definición 2.1.10. Un punto $x \in \mathbb{R}^n$ es el **valor adherente** a la sucesión $\{x_k\} \subseteq \mathbb{R}^n$ si y sólo si existe $\{x_{j_k}\} \subset \{x_k\}$ tal que $\lim_{k \rightarrow +\infty} x_{j_k} = x$.

Teorema 2.1.5. Sea $\{x_k\} \subseteq \mathbb{R}^n$. Son equivalentes:

1. $\{x_k\}$ es convergente.
2. $\{x_k\}$ es acotada y posee un único valor adherente.

Demostración. Véase Benazic (2000), página 23. □

2.1.3.4. Convergencia de sucesiones

Los resultados y definiciones sobre convergencia de sucesiones que presentamos en este apartado, serán usados para asegurar que la sucesión x^{k+1} generada tanto por el algoritmo MPPEI, como por sus dos variantes (MPPEI1 y MPPEI2), es acotada.

Lema 2.1.1. Sean $\{v_k\}$, $\{\lambda_k\}$ y $\{\beta_k\}$ sucesiones no negativas de números reales que satisfacen

$$v_k \leq (1 + \lambda_k)v_{k-1} + \beta_k,$$

donde $\sum_{k=1}^{+\infty} \beta_k < +\infty$, $\sum_{k=1}^{+\infty} \lambda_k < +\infty$. Entonces la sucesión $\{v_k\}$ es convergente.

Demostración. Véase Polyak (1987), Lema 2, página 44. □

Definición 2.1.11. Sea $\{x^k\} \subset \mathbb{R}^n$ tal que $\{x^k\}$ converge a un punto $\bar{x} \in \mathbb{R}^n$. Decimos que la convergencia es:

1. *lineal*, si existe una constante $\theta < 1$ y un número positivo $n \in \mathbb{N}$ tal que

$$\|x^k - \bar{x}\| \leq \theta \|x^{k-1} - \bar{x}\| \quad \forall k > n;$$

2. *superlineal*, si existe una sucesión $\{\beta_k\}$ convergente a cero y un número positivo $\bar{n} \in \mathbb{N}$ tal que

$$\|x^k - \bar{x}\| \leq \beta_k \|x^{k-1} - \bar{x}\| \quad \forall k > \bar{n}.$$

2.1.3.5. Series infinitas

En la presente no daremos una descripción detallada de las series infinitas, solo nos limitaremos a los conceptos básicos necesarios para los fines de ésta tesis. Para una descripción más detallada referirse a Espinoza (2004, cap. II) o Ayres y Mendelson (1991, cap. 47).

Definición 2.1.12. Sea $\{x_k\}_{k \geq 1}$ una sucesión de números reales. La **serie infinita** de números reales está denotada por:

$$\sum_{k=1}^{+\infty} x_k = x_1 + x_2 + \cdots + x_k + \cdots \quad (2.2)$$

donde x_k es llamado el k -ésimo término de la serie.

Observación 2.1.2. Con la serie infinita de números reales (2.2) formaremos una sucesión $\{s_k\}_{k \geq 1}$ definida de la siguiente forma:

$$\begin{aligned} s_1 &= x_1 \\ s_2 &= x_1 + x_2 \\ &\vdots \\ s_k &= x_1 + x_2 + \cdots + x_k \\ &\vdots \end{aligned}$$

A la sucesión $\{s_k\}_{k \geq 1}$ se le denomina sucesión de sumas parciales de la serie infinita (2.2), siendo s_k la k -ésima suma parcial de la serie.

Definición 2.1.13. Considerando la serie infinita (2.2) y la sucesión de sumas parciales $\{s_k\}$, decimos que:

1. Si $\lim_{k \rightarrow +\infty} s_k = s$ existe, entonces diremos que la serie infinita (2.2) es **convergente**, y converge hacia s , donde s es un número finito.

2. Si la serie infinita (2.2) es convergente, la **suma de la serie** se puede escribir como

$$\sum_{k=1}^{+\infty} x_k = \lim_{k \rightarrow +\infty} s_k = s.$$

Si la serie infinita es divergente, carece de suma.

Ejemplo 2.1.2. Hallar la suma de la serie infinita $\sum_{k=1}^{+\infty} \frac{1}{k(k+1)}$ en caso de ser convergente.

Solución. El k -ésimo término de la serie está dado por $x_k = \frac{1}{k(k+1)}$. Descomponiendo dicho término en fracciones parciales, se tiene que:

$$x_k = \frac{1}{k(k+1)} = \frac{A}{k} + \frac{B}{k+1},$$

efectuando las operaciones se tiene que $A = 1$ y $B = -1$, así que $x_k = \frac{1}{k} - \frac{1}{k+1}$. De esta forma:

$$\begin{aligned} x_1 &= 1 - \frac{1}{2} \\ x_2 &= \frac{1}{2} - \frac{1}{3} \\ &\vdots \\ x_k &= \frac{1}{k} - \frac{1}{k+1}. \end{aligned}$$

$$s_k = x_1 + x_2 + \cdots + x_k = 1 - \frac{1}{k+1}.$$

Por lo tanto $\lim_{k \rightarrow +\infty} s_k = 1$ existe, entonces la serie infinita $\sum_{k=1}^{+\infty} \frac{1}{k(k+1)}$ es convergente, y su suma es igual a 1, es decir $\sum_{k=1}^{+\infty} \frac{1}{k(k+1)} = 1$. \square

Proposición 2.1.5. Si $\sum_{k=1}^{+\infty} x_k$ es convergente, entonces $\lim_{k \rightarrow +\infty} x_k = 0$.

Demostración. Véase Espinoza (2004), página 103. \square

Proposición 2.1.6. Si $\lim_{k \rightarrow +\infty} x_k \neq 0$, entonces la serie infinita $\sum_{k=1}^{+\infty} x_k$ es divergente.

Demostración. Véase Espinoza (2004), página 104. \square

2.1.4. Conjuntos abiertos y cerrados

2.1.4.1. Interior de un conjunto

Definición 2.1.14. Sea $X \subseteq \mathbb{R}^n$:

1. Decimos que $x \in \mathbb{R}^n$ es un **punto interior** de X si y sólo si existe $\epsilon > 0$ tal que $B(x, \epsilon) \subseteq X$.

2. El conjunto de todos los puntos de interiores de X es llamado **interior** de X y será denotado por $\text{int}(X)$.
3. Decimos que $Y \subseteq \mathbb{R}^n$ es un **entorno o vecindad** de x si y sólo si $x \in \text{int}(Y)$.
4. Decimos que X es un conjunto abierto si y sólo si $X = \text{int}(X)$.

Observación 2.1.3. Para cualquier $X \subseteq \mathbb{R}^n$ se cumple que $\text{int}(X) \subseteq X$. En consecuencia, para probar que un conjunto es abierto, es suficiente probar que $X \subseteq \text{int}(X)$.

Teorema 2.1.6. Se cumplen las siguientes propiedades:

1. Los conjuntos \emptyset y \mathbb{R}^n son abiertos.
2. Si A_1 y A_2 son conjuntos abiertos, entonces $A_1 \cap A_2$ es un conjunto abierto.
3. Si $\{A_i\}_{i \in I}$ es una familia arbitraria de conjuntos abiertos, entonces $\cup_{i \in I} A_i$ es un conjunto abierto.

Demostración. Véase Benazic (2000), página 30. □

2.1.4.2. Clausura de un conjunto

Definición 2.1.15. Sea $X \subseteq \mathbb{R}^n$:

1. Decimos que $x \in \mathbb{R}^n$ es un **punto adherente** de X si y sólo si existe $\{x_n\} \subseteq X$ tal que $\lim_{n \rightarrow +\infty} x_n = x$.
2. El conjunto de todos los puntos de adherentes de X es llamado **clausura o cerradura** de X y será denotado por \overline{X} .
3. Decimos que X es un conjunto cerrado si y sólo si $X = \overline{X}$.

Teorema 2.1.7. Se cumplen las siguientes propiedades:

1. $X \subseteq \overline{X}$ para todo $X \subseteq \mathbb{R}^n$.
2. Si $X \subseteq Y$, entonces $\overline{X} \subseteq \overline{Y}$.

Demostración. Véase Benazic (2000), página 15. □

Teorema 2.1.8. Sean $X \subseteq \mathbb{R}^n$ y $x \in \mathbb{R}^n$. Son equivalentes:

1. $x \in \overline{X}$.
2. $B(x, \epsilon) \cap X \neq \emptyset$, para todo $\epsilon > 0$.

Demostración. Véase Benazic (2000), página 36. □

Corolario 2.1.2. $x \notin \overline{X}$ si y sólo si existe $\epsilon > 0$ tal que $B(x, \epsilon) \subseteq \mathbb{R}^n - X$.

Demostración. Véase Benazic (2000), página 36. \square

Teorema 2.1.9. Sea $X \subseteq \mathbb{R}^n$. X es cerrado si y sólo si su complemento $X^c = \mathbb{R}^n - X$ es abierto.

Demostración. Véase Benazic (2000), página 37. \square

Teorema 2.1.10. Se cumplen las siguientes propiedades:

1. Los conjuntos \emptyset y \mathbb{R}^n son cerrados.
2. Si A_1 y A_2 son conjuntos cerrados, entonces $A_1 \cup A_2$ es un conjunto cerrado.
3. Si $\{A_i\}_{i \in I}$ es una familia arbitraria de conjuntos cerrados, entonces $\bigcap_{i \in I} A_i$ es un conjunto cerrado.

Demostración. Véase Benazic (2000), página 37. \square

2.1.4.3. Puntos de acumulación

Definición 2.1.16. Sea $X \subseteq \mathbb{R}^n$:

1. Decimos que $x \in \mathbb{R}^n$ es un **punto de acumulación** o punto límite de X si y sólo si $B(x, \epsilon) \cap (X - \{x\}) \neq \emptyset$ para todo $\epsilon > 0$.
2. El conjunto de todos los puntos de acumulación de X es llamado **conjunto derivado** de X y será denotado por X' . Por lo tanto, $x \in X'$ si y sólo si $x \in \overline{X - \{x\}}$.
3. Si $x \in X$ no es un punto de acumulación de X , entonces decimos que x es un punto aislado de X , esto significa que existe $\epsilon > 0$, tal que x es el único punto de X en $B(x, \epsilon)$.
4. Decimos que X es un **conjunto discreto** si y sólo si todos sus puntos son aislados.

En vista de la definición, es evidente que todos los puntos de acumulación son puntos de adherencia, así que en general $X' \subseteq \overline{X}$, sin embargo el recíproco por lo general no es cierto. La estrecha relación entre dichos puntos, se pone de manifiesto en el siguiente resultado:

Proposición 2.1.7. Sea $X \subseteq \mathbb{R}^n$. $x \in X'$ si y sólo si $x \in \overline{X - \{x\}}$.

Demostración. Véase Benazic (2000), página 40. \square

Teorema 2.1.11. Sean $X \subseteq \mathbb{R}^n$ y $x \in \mathbb{R}^n$. Las siguientes afirmaciones son equivalentes.

1. $x \in X'$.
2. Existe una sucesión $\{x_n\} \subseteq X - \{x\}$ tal que $\lim_{n \rightarrow +\infty} x_n = x$.

Demostración. Véase Benazic (2000), página 41. □

Teorema 2.1.12. (Weierstras) Todo conjunto infinito y acotado $X \subseteq \mathbb{R}^n$ tiene al menos un punto de acumulación.

Demostración. Véase Benazic (2000), página 41. □

2.1.5. Funciones continuas

Definición 2.1.17. Sean $X \subset \mathbb{R}^m$ y $f : X \rightarrow \mathbb{R}^n$.

1. Decimos que f es continua en $\bar{x} \in X$, si y sólo si para todo $\epsilon > 0$, existe un $\delta = \delta(\epsilon) > 0$ tal que si $x \in X$ y $\|x - \bar{x}\| < \delta$, entonces $\|f(x) - f(\bar{x})\| < \epsilon$.
2. Decimos que f es discontinua en $\bar{x} \in X$, si y sólo si f no es continua en \bar{x} .
3. Decimos que f es continua en $Y \subseteq X$, si y sólo si f es continua en y , para todo $y \in Y$.

Teorema 2.1.13. Sean $F : \mathbb{R}^m \supseteq X \rightarrow \mathbb{R}^n$ y $x \in X$. Son equivalentes:

1. F es continua en x .
2. $\{x_k\} \subseteq X$ tal que $\lim_{k \rightarrow +\infty} x_k = x$, entonces $\lim_{k \rightarrow +\infty} F(x_k) = F(x)$.

Demostración. Véase Benazic (2001), página 38. □

Corolario 2.1.3. Sean $f, g : \mathbb{R}^m \supseteq X \rightarrow \mathbb{R}^n$ y $\alpha : \mathbb{R}^m \supseteq X \rightarrow \mathbb{R}$ funciones continuas en $x \in X$, entonces $f \pm g$, αf , $\|f\|$, $\langle f, g \rangle$, $d(f, g)$ y $\frac{1}{\alpha}$ (con $\alpha(x) \neq 0$) son funciones continuas en x .

Demostración. Véase Benazic (2001), página 39. □

Teorema 2.1.14. Sea $f : \mathbb{R}^m \supseteq X \rightarrow \mathbb{R}^n$ una función vectorial, donde $f = (f_1, \dots, f_n)$, y $x \in X$. Son equivalentes:

1. f es continua en x .
2. f_i son continuas en x , para todo $i \in \{1, \dots, n\}$.

Demostración. Véase Benazic (2001), página 39. □

2.1.6. Funciones de Lipschitz

Definición 2.1.18. Sean $X \subseteq \mathbb{R}^m$ y $f : X \rightarrow \mathbb{R}^n$. Decimos que f es **Lipschitziana** en X , si y sólo si existe $K > 0$ tal que:

$$|f(x) - f(y)| \leq K \|x - y\| \quad \forall x, y \in X.$$

Ejemplo 2.1.3. La función norma $f : \mathbb{R}^n \rightarrow \mathbb{R}$ definida por $f(x) = \|x\|$ es Lipschitziana. En efecto

$$|f(x) - f(y)| = ||x| - |y|| \leq \|x - y\| \quad \forall x, y \in X.$$

Proposición 2.1.8. Sea $X \subseteq \mathbb{R}^m$ y $f : X \rightarrow \mathbb{R}^n$. Si f es Lipschitziana en X , entonces f es continua en X .

Demostración. Véase Benazic (2001), página 41. □

2.2. Elementos básicos del análisis convexo

2.2.1. Generalidades de las funciones de \mathbb{R}^n en $\mathbb{R} \cup \{+\infty\}$

Definición 2.2.1. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ una función de valores extendidos.

1. El **dominio efectivo** de f , denotado por $\text{dom}(f)$, está dado por

$$\text{dom}(f) = \{x \in \mathbb{R}^n : f(x) < +\infty\}.$$

2. Decimos que f es **propia**, si:

- a) $\text{dom}(f) \neq \emptyset$,
- b) para todo $x \in \text{dom}(f)$, $f(x) > -\infty$.

Observación 2.2.1. Toda función $g : \mathbb{R}^n \supseteq X \rightarrow \mathbb{R}$, siempre puede ser prolongada a una función de valores extendidos¹ $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ definida por:

$$f(x) = \begin{cases} g(x) & \text{si } x \in X \\ +\infty & \text{si } x \notin X. \end{cases}$$

Definición 2.2.2. Sean $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ y $\alpha \in \mathbb{R}$.

1. El **conjunto de nivel** (o conjunto de nivel inferior) de la función f está definido por

$$L_f(\alpha) = \{x \in \mathbb{R}^n : f(x) \leq \alpha\}.$$

¹En general, puede ser extendida a una función $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, donde $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$.

2. El **epígrafo** de la función f está definido por el siguiente subconjunto de $\mathbb{R}^n \times \mathbb{R}$

$$\text{epi}(f) = \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R} : f(x) \leq \alpha\}.$$

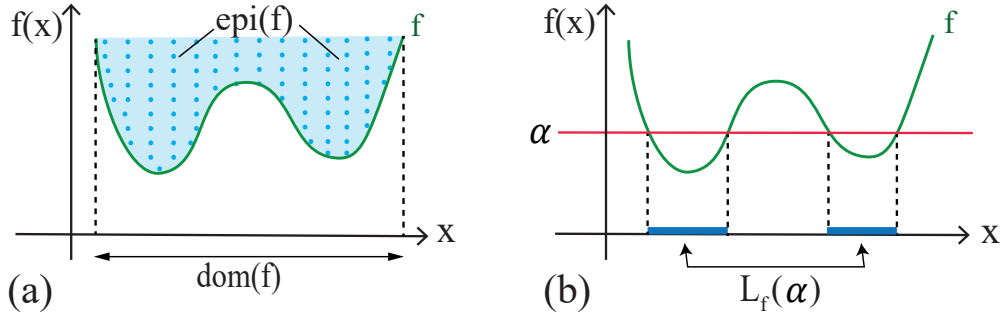


Figura 2.1: (a) Epígrafo y (b) conjunto de nivel inferior de una función no convexa.

Definición 2.2.3. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ una función propia. Decimos que f es **localmente Lipschitz** en $x \in \text{dom}(f)$, si existe algún $\epsilon_x > 0$ tal que:

$$|f(z) - f(y)| \leq L_x \|z - y\| \quad \forall z, y \in B(x, \epsilon_x),$$

donde $B(x, \epsilon_x) = \{y \in \mathbb{R}^n, \|y - x\| < \epsilon_x\}$ y $L_x > 0$ es la constante de Lipschitz. Decimos que f es localmente Lipschitz en \mathbb{R}^n , si f es localmente Lipschitz en cada $x \in \mathbb{R}^n$.

Definición 2.2.4. Sean $X \subseteq \mathbb{R}^n$ y $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ una función.

1. Decimos que una sucesión $\{x^k\} \subseteq X$ es **crítica** (en relación al conjunto X), si $\lim_{k \rightarrow +\infty} \|x^k\| = +\infty$ o $\lim_{k \rightarrow +\infty} x^k = \bar{x} \in \overline{X} \setminus X$, donde la notación $\bar{x} \in \overline{X} \setminus X$ significa que $\bar{x} \in \overline{X}$ y $\bar{x} \notin X$.
2. Decimos que f es **coerciva** en X , si para toda sucesión crítica $\{x^k\} \subseteq X$ se tiene que

$$\limsup_{k \rightarrow +\infty} f(x^k) = +\infty,$$

o equivalente

$$\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty.$$

Ejemplo 2.2.1. Sea $x^k \in \mathbb{R}^n$ un punto fijo. La función $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, definida por $f(\cdot) = \|\cdot - x^k\|^2$, es coerciva.

En efecto, por la Proposición 2.1.1 se tiene que

$$\| \|x\| - \|x^k\| \|^2 \leq \|x - x^k\|^2 = f(x) \quad \forall x, x^k \in \mathbb{R}^n.$$

Dado que $\|x\| \rightarrow +\infty$ y $\|x^k\|$ representa un número fijo, al reemplazar dichos valores en la expresión anterior fácilmente se obtiene que $f(x) \rightarrow +\infty$. \square

Definición 2.2.5. Sea $f : \mathbb{R}^n \supseteq X \rightarrow \mathbb{R} \cup \{+\infty\}$ una función.

1. f es **semicontinua inferior** (sci) en $\bar{x} \in X$, si para toda sucesión $\{x^k\}$ de X convergente a \bar{x} , se tiene que

$$f(\bar{x}) \leq \liminf_{k \rightarrow +\infty} f(x^k) := \sup_{n \in \mathbb{N}} \inf_{k \geq n} \{f(x^k)\},$$

2. f es **semicontinua superior** (scs) en $\bar{x} \in \mathbb{R}^n$, si para toda sucesión $\{x^k\}$ de X convergente a \bar{x} , se tiene que

$$\limsup_{k \rightarrow +\infty} f(x^k) \leq f(\bar{x}).$$

Si f es sci (scs) para todo $x \in X \subseteq \mathbb{R}^n$, entonces diremos que f es sci (scs) en X , o simplemente que f es sci (scs).

De la definición resulta claro que toda función continua es semicontinua inferior, sin embargo, como se muestra en el siguiente ejemplo, el recíproco no es verdad.

Ejemplo 2.2.2. (Papa Quiroz, 2017, Ejemplo 3.1, p. 35). La función $f : [1, 2] \rightarrow \mathbb{R}$ definida por

$$f(x) = \begin{cases} x^2 + 1 & \text{si } x \in (1, 2], \\ 1 & \text{si } x = 1, \end{cases}$$

no es continua en $\hat{x} = 1$, sin embargo, como se prueba a continuación, es sci en dicho punto. En efecto, sea $\{x^k\} \subseteq [1, 2]$ tal que $x^k \rightarrow 1$, como f es creciente, para $n \in \mathbb{N}$ tenemos que $\inf_{k \geq n} f(x^k) = 2$, por lo tanto $\sup_{n \in \mathbb{N}} \inf_{k \geq n} f(x^k) = 2 > f(\bar{x}) = 1$.

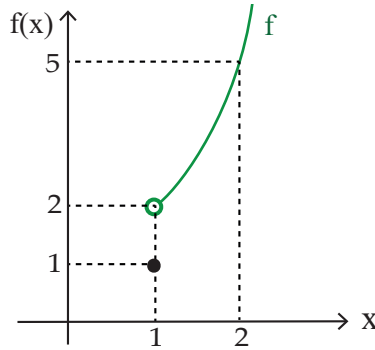


Figura 2.2: La función f es sci, sin embargo no es continua.

La importancia natural de la sci en el estudio de las funciones se evidencia por ejemplo en el siguiente resultado.

Teorema 2.2.1. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ una función. Son equivalentes:

1. f es sci,
2. $\text{epi}(f)$ es cerrado,

3. $L_f(\alpha)$ es cerrado para todo $\alpha \in \mathbb{R}$.

Demostración. Véase Crouzeix et al. (2003), página 16. \square

Definición 2.2.6. Sea $C \subset \mathbb{R}^n$, $C \neq \emptyset$, definimos la **función indicatriz** de C como

$$\delta_C(x) = \begin{cases} 0 & \text{si } x \in C \\ +\infty & \text{si } x \notin C. \end{cases}$$

Definición 2.2.7. Sean $C \subset \mathbb{R}^n$ y $\bar{x} \in C$. El **cono normal** al punto \bar{x} relacionado al conjunto C está dado por

$$N_C(\bar{x}) = \{d \in \mathbb{R}^n : \langle d, x - \bar{x} \rangle \leq 0, \quad \forall x \in C\}.$$

Proposición 2.2.1. Sea $C \subset \mathbb{R}^n$, $C \neq \emptyset$. El conjunto C es cerrado si y sólo si $\delta_C(\cdot)$ es sci.

Demostración. Véase Crouzeix et al. (2003), página 38. \square

Proposición 2.2.2. Si $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ es una función propia, semicontinua inferior y coerciva, entonces el valor óptimo f^* es finito, y además el conjunto $\arg \min\{f(x) : x \in \mathbb{R}^n\}$ es no vacío y compacto.

Demostración. Véase Rockafellar y Wets (2009), Teorema 1.9. \square

2.2.2. Convexidad y cuasi-convexidad de funciones

En la presente subsección presentamos algunas definiciones básicas, necesarios para los fines de esta tesis, relacionadas a la convexidad y cuasi-convexidad de funciones del tipo $f : \mathbb{R}^n \supseteq X \rightarrow \mathbb{R}$, las cuáles en función de la Observación 2.2.1, fácilmente pueden ser generalizadas a funciones con valores en el conjunto de los reales extendidos $f : \mathbb{R}^n \supseteq X \rightarrow \mathbb{R} \cup \{+\infty\}$. Dichas definiciones han sido tomadas de Van Tiel (1984) y Berkovitz (2003), por lo que se recomienda revisar dichas fuentes para mayor detalle.

Definición 2.2.8. Un conjunto $X \subseteq \mathbb{R}^n$ es llamado **convexo** si para todo $x, y \in X$ y para todo $\alpha \in [0, 1]$ se cumple que:

$$\alpha x + (1 - \alpha)y \in X.$$

Definición 2.2.9. Sea $X \subseteq \mathbb{R}^n$ un subconjunto convexo no vacío. Una función $f : X \rightarrow \mathbb{R}$ es llamada **convexa** en X , si para todo $x, y \in X$ y para todo $\alpha \in [0, 1]$ se cumple que:

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y).$$

Si la desigualdad es estricta para cada $\alpha \in (0, 1)$ entonces se dice que f es estrictamente convexa.

Definición 2.2.10. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función. Decimos que f es **cuasi-convexa** en \mathbb{R}^n si para todo $x, y \in \mathbb{R}^n$ y para todo $\alpha \in [0, 1]$ se cumple que

$$f(\alpha x + (1 - \alpha)y) \leq \max\{f(x), f(y)\}.$$

Observación 2.2.2. 1. De las definiciones precedentes, es inmediato ver que la convexidad implica cuasi-convexidad, sin embargo lo recíproco no siempre es verdadero. Por ejemplo considere la función $f(x, y) = -e^{-x^2-y^2}$ es cuasi-convexa pero no es convexa.

2. Mientras que una función convexa puede ser caracterizada por la convexidad de su epígrafo, una función cuasi-convexa puede ser caracterizada por la convexidad de su conjunto de nivel.

Ejemplo 2.2.3. Son cuasiconvexas las siguientes funciones:

1. $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^3$, $f(x) = \arctg(x)$;
2. $f : \mathbb{R}_{>} \rightarrow \mathbb{R}$, $f(x) = \log(x)$;
3. $f : \mathbb{R}^2 \rightarrow \mathbb{R} \cup \{+\infty\}$, tal que

$$f(x, y) = \begin{cases} 0 & \text{si } x < 0 \text{ o } x = 0, y \leq 0, \\ \sqrt{y} & \text{si } x > 0, y \geq 0, \\ +\infty & \text{caso contrario.} \end{cases}$$

Definición 2.2.11. Sea $F : \mathbb{R}^n \rightarrow \mathbb{R}^p$ una función. Decimos que F es \mathbb{R}_{\leq}^p -cuasi-convexa si y sólo si cada componente de la función F ($F_i : \mathbb{R}^n \rightarrow \mathbb{R}$) es cuasi-convexa.

Definición 2.2.12. Sea $F : \mathbb{R}^n \rightarrow \mathbb{R}^p$ una función. Decimos que F es \mathbb{R}_{\leq}^p -convexa si cada componente de la función F ($F_i : \mathbb{R}^n \rightarrow \mathbb{R}$) es convexa.

Definición 2.2.13. Sea $F : \mathbb{R}^n \rightarrow \mathbb{R}^p$ una función. Decimos que F es **localmente Lipschitz** en \mathbb{R}^n si cada componente de la función F ($F_i : \mathbb{R}^n \rightarrow \mathbb{R}$) es localmente Lipschitz en \mathbb{R}^n .

2.2.3. Derivadas Direccionales, Subgradientes y Subdiferenciales

En análisis la herramienta fundamental para obtener condiciones necesarias de optimalidad, es la célebre Regla de Fermat, que en el caso diferenciable enuncia que la derivada se anula en un punto de mínimo o de máximo local. Esta regla, válida en el contexto de la optimización convexa ², puede ser generalizada en términos del vector gradiente asociado con un hiperplano tangente al gráfico de la función; y es válida en diversos contextos que van desde la programación matemática, el cálculo variacional, hasta la teoría de control.

Cuando la diferenciabilidad en el sentido clásico falla (en particular cuando las funciones no necesariamente son convexas), es natural preguntarse si es posible extender la noción de derivada, de modo que se puedan recuperar al menos algunas de sus propiedades. La respuesta a tal interrogante es afirmativa, y está relacionada con la derivada lateral a la derecha (de una función convexa no necesariamente diferenciable), la cuál puede ser escrita en términos de vectores **subgradientes** asociados con hiperplanos soportes al epigrafo de la función. Es así que se concibió la noción de subdiferencial. En lo sucesivo nos referiremos a la derivada direccional derecha como **derivada direccional**.

Los conceptos y resultados que presentamos a continuación, fueron tomados de Bazaraa et al. (2013), Rockafellar (1970) y Navarro (2013).

Definición 2.2.14. Sean $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ una función convexa propia, $x \in \text{dom}(f)$ y $d \in \mathbb{R}^n$. La **derivada direccional** de f en x en la dirección $d \in \mathbb{R}^n$ es definida (en el sentido usual) como:

$$f'(x, d) = \lim_{t \rightarrow 0^+} \frac{f(x + td) - f(x)}{t},$$

si este límite existe ($+\infty$ es permitido como límite).

Si f es diferenciable en x , entonces la derivada direccional existe en toda dirección $d \in \mathbb{R}^n$, $f'(x, d)$ es una función lineal en d y tenemos la siguiente relación:

$$f'(x, d) = \langle \nabla f(x), d \rangle. \quad (2.3)$$

Definición 2.2.15. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ una función convexa ³.

1. Decimos que $d \in \mathbb{R}^n$ es el **subgradiente** de f en $x \in \mathbb{R}^n$ (en el sentido del análisis convexo), si cumple la siguiente desigualdad:

$$f(y) \geq f(x) + \langle d, y - x \rangle \quad \forall y \in \mathbb{R}^n.$$

²Las funciones convexas poseen algunas propiedades usuales de diferenciabilidad, y una de ellas es la existencia de la derivada direccional lateral (derecha e izquierda) en toda dirección en un punto interior de su dominio.

³La definición tiene sentido para funciones más generales, en las cuáles se mantienen algunas de sus propiedades. Sin embargo para los propósitos de esta tesis, consideraremos sólo funciones convexas

2. El **subdiferencial** de f en x , denotado por $\partial f(x)$, es el conjunto de todos los subgradiientes de f en x , esto es

$$\partial f(x) = \{d \in \mathbb{R}^n : f(y) \geq f(x) + \langle d, y - x \rangle, \quad \forall y \in \mathbb{R}^n\}.$$

Si $\partial f(x) \neq \emptyset$ decimos que f es subdiferenciable en x .

Geométricamente, $d \in \mathbb{R}^n$ es un subgradiente de la función convexa f en x , si el gráfico de f en \mathbb{R}^{n+1} esta por encima del gráfico del hiperplano $h = f(x) + \langle d, y - x \rangle$. Como $(x, f(x))$ esta en el hiperplano h , este hiperplano constituye un hiperplano soporte a $\text{epi}(f)$ en $(x, f(x))$. Esto es, la existencia de un vector subgradiente establece la existencia de un hiperplano soporte no vertical a $\text{epi}(f)$ en $(x, f(x))$.

En la Figura (2.3), en la cual mostramos un ejemplo para el caso $n = 1$, podemos ver que el subgradiente d_1 define un hiperplano $h_1(x_1) = f(x_0) + \langle d_1, x_1 - x_0 \rangle$ que soporta al conjunto $\text{epi}(f)$ en el punto frontera $(x_0, f(x_0))$.

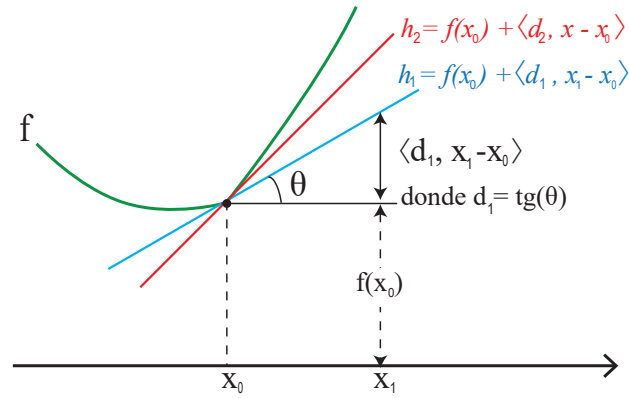


Figura 2.3: d_1 y d_2 son los subgradiientes de f en x_0 .

Ejemplo 2.2.4. Cero es un subgradiente de la función $f(x) = \sin(x)$ en el punto $x = -\frac{\pi}{2}$, dado que $\sin(x) \geq -1 + 0 \cdot (x + \frac{\pi}{2})$, para toda $x \in \mathbb{R}$.

Ejemplo 2.2.5. Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ definida por $f(x) = |x|$ (ver Figura 2.4), entonces

$$\partial f(x) = \begin{cases} [-1, 1] & \text{si } x = 0, \\ 1 & \text{si } x > 0, \\ -1 & \text{si } x < 0. \end{cases}$$

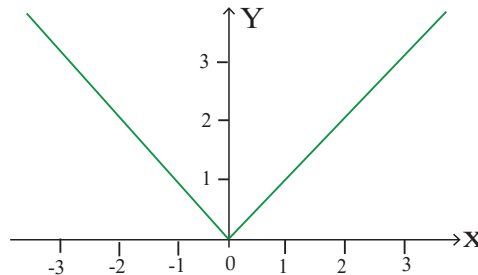


Figura 2.4: La función valor absoluto.

En efecto, consideremos los siguientes casos:

i) Sea $x = 0$, entonces

$$\partial f(0) = \{d \in \mathbb{R} : |y| \geq |0| + d.y, \forall y \in \mathbb{R}\}.$$

Observamos que:

- Si $y = 0$, entonces d puede tomar cualquier valor real, esto es $d \in \mathbb{R}$.
- Si $y > 0$, entonces $y \geq d.y$, lo cual implica que $d \in (-\infty, 1]$.
- Si $y < 0$, entonces $-y \geq d.y$, así que $d \in [-1, +\infty)$.

Interceptando los conjuntos tenemos que $d \in [-1, 1]$; así que $\partial f(0) = [-1, 1]$.

ii) Sea $x > 0$, entonces

$$\partial f(x) = \{d \in \mathbb{R} : |y| \geq x + d(y - x), \forall y \in \mathbb{R}\}.$$

- Si $y = 0$, entonces $0 \geq x - d.x$, lo que implica que $d \in [1, +\infty)$.
- Si $y > 0$ y $y \neq x$, tenemos que $y - x > 0$ ó $y - x < 0$.
 - Si $y - x > 0$ tenemos que $y - x \geq d(y - x)$, lo que implica que $d \in (-\infty, 1]$.
 - Si $y - x < 0$ tenemos que $y - x \geq d(y - x)$, así que $d \in [1, +\infty)$.

De ambos casos obtenemos que $d = 1$.

- Si $y > 0$ y $y = x$, entonces la desigualdad $y - x \geq d(y - x)$ se cumple trivialmente para cualquier $d \in \mathbb{R}$. En particular para $d = 1$.
- Finalmente si $y < 0$, se tiene que $-y \geq x + d(y - x)$, lo cuál es verdad para $d = 1$.

Por lo tanto $\partial f(x) = \{1\}$, $\forall x > 0$.

iii) Para $x < 0$, se procede de manera similar al caso del item anterior, obteniendo $\partial f(x) = \{-1\}$, $\forall x < 0$.

Teorema 2.2.2. Una función convexa $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es diferenciable, si y solo si el conjunto $\partial f(x)$ contiene un único elemento, se denotara como $\partial f(x) = \{\nabla f(x)\}$.

Demostración. Véase Bazaraa et al. (2013), Lema 3.3.2, página 110. \square

Teorema 2.2.3. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ convexa con constante de Lipschitz $K > 0$ en $x \in \mathbb{R}^n$, entonces:

1. La derivada direccional en cada dirección $d \in \mathbb{R}^n$ existe y satisface

$$f'(x, d) = \inf_{t>0} \frac{f(x + td) - f(x)}{t};$$

2. La función $y \rightarrow f'(x, d)$ es positiva homogénea y subaditiva en \mathbb{R}^n con

$$f'(x, d) \leq K\|d\|;$$

3. $f'(x, d)$ es semicontinua superior como una función de (x, d) y Lipschitz con constante K como función de d .

4. $-f'(x, d) \leq f'(x, d)$.

Demostración. Véase Navarro (2013), Teorema 2.1, página 43. \square

Teorema 2.2.4. Sean $f : \mathbb{R}^n \rightarrow \mathbb{R}$ convexa y $x \in \mathbb{R}^n$ (fijo, arbitrario). Se cumple:

1. $\partial f(x) = \{w \in \mathbb{R}^n : \langle w, d \rangle \leq f'(x, d), \quad \forall d \in \mathbb{R}^n\}$

2. $\partial f(x)$ es no vacío, convexo y compacto y $\partial f(x) \subseteq B(x, K)$, donde K es la constante de Lipschitz de f en x .

3. $f'(x, d) = \max\{\langle w, d \rangle : w \in \partial f(x), \forall d \in \mathbb{R}^n\}$.

Demostración. Véase Navarro (2013), Teorema 2.2, página 47. \square

2.3. Compacidad y existencia de minimizadores

En la presente sección estudiamos las condiciones suficientes para garantizar la existencia de soluciones de problemas de optimización, las cuáles, para un mejor entendimiento del tema, irán acompañadas de ejemplos ilustrativos. Cabe resaltar que dicha información ha sido tomada de Papa Quiroz (2017), por lo que se recomienda revisar dicha fuente, para mayores detalles.

2.3.1. Óptimos locales y globales

Definición 2.3.1. Sean $f : \mathbb{R}^n \supseteq X \rightarrow \mathbb{R}$ una función y $\bar{x} \in X$. Decimos que

1. \bar{x} es un punto de mínimo local de f sobre X si existe un $\epsilon > 0$ tal que

$$f(\bar{x}) \leq f(x) \quad \forall x \in X \cap B(\bar{x}, \epsilon). \quad (2.4)$$

2. \bar{x} es un punto de mínimo global de f sobre X si

$$f(\bar{x}) \leq f(x) \quad \forall x \in X. \quad (2.5)$$

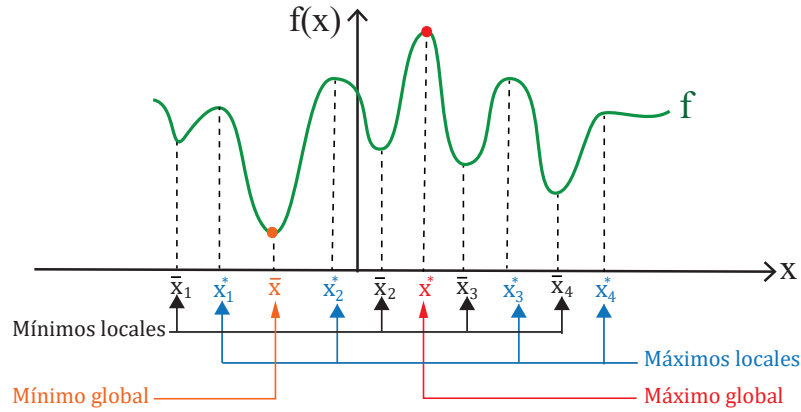


Figura 2.5: Óptimos locales y globales.

- Observación 2.3.1.**
1. *Invirtiendo las desigualdades de las relaciones (2.4) y (2.5), obtenemos las definiciones de punto de máximo local y máximo global, respectivamente.*
 2. *La relación inmediata de estas dos definiciones es que todo mínimo (máximo) global es un mínimo (máximo) local, sin embargo el recíproco evidentemente no es verdad.*
 3. *Todo problema de maximización puede ser transformado a uno de minimización, o viceversa, anteponiendo el signo negativo; por ejemplo maximizar f es equivalente a minimizar $-f$.*

2.3.2. Resultados de existencia

Consideremos el problema

$$\min_{x \in X} f(x) = \min\{f(x) : x \in X\},$$

donde $f : \mathbb{R}^n \supseteq X \rightarrow \mathbb{R}$. Una pregunta natural es saber si existe un punto que resuelva el problema planteado. Veremos que para ello será necesario imponer algunas condiciones sobre f y X .

Una condición necesaria para probar la existencia de un punto de mínimo global es que la función f sea acotada inferiormente, esto es, $\inf\{f(x) : x \in X\} > -\infty$. Lamentablemente esta condición no es suficiente, considere por ejemplo la función exponencial $f(x) = e^x$, donde $x \in \mathbb{R}$. Esta función se encuentra acotada inferiormente por cero, sin embargo, como se puede observar en la Figura 2.6, no existe un punto de mínimo global.

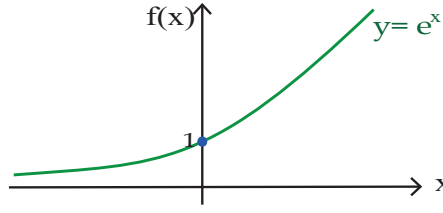


Figura 2.6: La función exponencial tiene ínfimo pero no mínimo.

Un resultado clásico para garantizar la existencia de puntos de mínimo global es el conocido Teorema de Bolzano-Weierstrass, aplicado a funciones continuas sobre conjuntos compactos. En esta tesis consideraremos una hipótesis más débil, la condición de semicontinuidad inferior (ver Definición 2.2.5 (1) - página 17).

Teorema 2.3.1. Si $f : \mathbb{R}^n \supseteq X \rightarrow \mathbb{R}$ es una función semicontinua inferior en un conjunto no vacío y compacto X , entonces f logra su mínimo global sobre X .

Demostración. Véase Papa Quiroz (2017), Teorema 3.1. □

Ejemplo 2.3.1. Consideremos la función definida en el Ejemplo 2.2.2, como f es semicontinua inferior y $[1, 2]$ es un conjunto compacto, entonces por el teorema anterior concluimos que f admite un mínimo global en $[1, 2]$. □

Corolario 2.3.1. Dada una función $f : \mathbb{R}^n \supseteq X \rightarrow \mathbb{R}$. Si X es un conjunto no vacío y compacto y f es continua en X , entonces existe un punto de mínimo global.

Demostración. Véase Papa Quiroz (2017), Corolario 3.1. □

Observación 2.3.2. La condición de compacidad de X es muy fuerte para garantizar la existencia de un mínimo global. En efecto, considere el problema

$$\min\{x^2 : x \in \mathbb{R}\}. \quad (2.6)$$

Aquí $X = \mathbb{R}$ no es compacto (pues no cumple el requisito de acotación), sin embargo existe el mínimo global.

Para debilitar la hipótesis de compacidad de X , en el siguiente resultado incorporamos el concepto de **conjunto de nivel inferior** de f (ver Definición 2.2.2 (1) - página 15).

Corolario 2.3.2. Sean $f : \mathbb{R}^n \supseteq X \rightarrow \mathbb{R} \cup \{+\infty\}$ una función y $\alpha \in \mathbb{R}$. Si f es semicontinua inferior en un conjunto no vacío y compacto $L_f(\alpha)$, entonces existe un punto de mínimo global de f en X .

Demostración. Véase Papa Quiroz (2017), Corolario 3.2. □

Para debilitar más aún las condiciones de existencia de mínimos globales en el siguiente resultado incorporamos el concepto de **función coerciva** (ver Definición (2) - página 26).

Teorema 2.3.2. *Si una función $f : \mathbb{R}^n \supseteq X \rightarrow \mathbb{R} \cup \{+\infty\}$ es coerciva y semi-continua inferior en un conjunto $X \neq \emptyset$, entonces existe un punto que representa un mínimo global de f en X .*

Demostración. Véase Papa Quiroz (2017), Corolario 3.4. □

Capítulo 3

SUBDIFERENCIALES GENERALIZADOS Y DIRECCIONES DE DESCENSO

3.1. Subdiferenciales Generalizados

La teoría moderna de optimización tubo sus inicios esencialmente con el desarrollo del método simplex, formulado por George Dantzig en 1947¹; sin embargo, los enfoques modernos de la teoría de optimización deben sus orígenes al cálculo de variaciones que ha sido estudiado por más de tres siglos, y que también fue crucial en el desarrollo del análisis funcional.

*El estudio de la optimización convexa², actualmente conocida como **análisis convexo** (Termino que fue sugerido por Albert. W. Tucker, famoso por las condiciones de Kuhn-Tucker), se inició con la publicación de la obra ya clásica “Convex Analysis”(Rockafellar, 1970), la cual tubo muchísimo impacto en el desarrollo de la teoría y los métodos de optimización; en cuyo proceso, los investigadores encontraron que muchos de los problemas de optimización convexos no eran diferenciables en el punto mínimo; es así que las funciones no diferenciables, marcaron un paradigma en la teoría moderna de optimización, el cuál condujo al desarrollo de un enfoque completamente diferente, concibiéndose de esta manera la noción de **subdiferencial**.*

El subdiferencial de una función en un punto dado, es un conjunto que puede ser considerado como un sustituto de la noción de la derivada en puntos donde la función no es diferenciable. Gracias al estudio del mismo, se pudieron desarrollar un conjunto de reglas muy útiles para llevar a cabo el análisis de funciones no diferenciables. Es así que de una manera natural se logró extender el dominio de la optimización no diferenciable más allá de la convexidad. La primera extensión natural se dio para funciones localmente Lipschitz, lo cuál fue iniciado por Clarke (1990), quien fue estudiante de Rockafellar. Él introdujo la noción de gradiente generalizado para funciones localmente Lipschitz. Es así que a partir de

¹dicho modelo matemático (formulado durante la segunda guerra mundial) tenía como objetivo reducir los costos del ejército y aumentar las pérdidas del enemigo

²Las funciones convexas poseen algunas propiedades usuales de diferenciabilidad, y una de ellas es la existencia de la derivada direccional lateral (derecha e izquierda) en toda dirección en un punto interior de su dominio.

la publicación de Rockafellar (1970), Clarke (1990), entre otras, el uso de subdiferenciales (como el subdiferencial Fréchet, y el de Clarke) se ha convertido en algo tan cotidiano como el uso de gradientes.

En esta sección presentamos algunos conceptos y resultados relacionados a los subdiferenciales generalizados que se utilizarán más adelante, los cuales fueron tomados de Clarke (1990), Clarke (2013) y Rockafellar y Wets (2009, cap. 8 y 10). Si bien la mayoría de los conceptos que siguen pueden ser desarrollados para funciones del tipo $f : X \rightarrow Y$, donde X, Y son espacios normados, espacios vectoriales topológicos localmente convexos, o más generalmente sobre espacios de Banach, en la presente solamente nos limitaremos a los espacios Euclidianos, y a funciones del tipo $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$.

3.1.1. Subdiferencial de Fréchet y en el límite

Definición 3.1.1. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ una función propia.

1. Para cada $x \in \text{dom}(f)$, donde $f(x)$ es finito, el conjunto de subgradien-
tes regulares (también llamado **subdiferencial de Fréchet**) de f en x , de-
notado por $\hat{\partial}f(x)$ esta formado por un conjunto de vectores $d \in \mathbb{R}^n$, tales
que

$$f(y) \geq f(x) + \langle d, y - x \rangle + \theta(\|y - x\|), \text{ donde } \lim_{\substack{y \rightarrow x \\ y \neq x}} \frac{\theta(\|y - x\|)}{\|y - x\|} = 0,$$

o equivalentemente

$$\hat{\partial}f(x) = \left\{ d \in \mathbb{R}^n : \liminf_{\substack{y \rightarrow x \\ y \neq x}} \frac{f(y) - f(x) - \langle d, y - x \rangle}{\|y - x\|} \geq 0 \right\}.$$

Si $x \notin \text{dom}(f)$ entonces $\hat{\partial}f(x) = \emptyset$.

2. El conjunto de subgradien-
tes generalizados (también llamado **subdiferen-
cial en el límite**) de f en x , denotado por $\partial^{Lim}f(x)$ esta definido por

$$\partial^{Lim}f(x) = \left\{ d \in \mathbb{R}^n : \exists x^l \rightarrow x, f(x^l) \rightarrow f(x), d^l \in \hat{\partial}f(x^l) \text{ y } d^l \rightarrow d \right\}.$$

Proposición 3.1.1. Para una función $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ y un punto $\bar{x} \in \text{dom}(f)$ en el cuál f es finito, los conjuntos subgradien-
tes $\partial^{Lim}f(\bar{x})$ y $\hat{\partial}f(\bar{x})$ son
cerrados, con $\hat{\partial}f(\bar{x})$ convexo y $\hat{\partial}f(\bar{x}) \subset \partial^{Lim}f(\bar{x})$.

Demostración. (a) Probaremos que $\hat{\partial}f(\bar{x})$ es convexo. En efecto, sean $d^1, d^2 \in \hat{\partial}f(\bar{x})$, entonces para todo $\alpha \in [0, 1]$ se tiene que

$$\begin{aligned} \alpha f(y) &\geq \alpha f(\bar{x}) + \langle \alpha d^1, y - \bar{x} \rangle + \alpha \theta_1(\|y - \bar{x}\|) \quad \forall y \in \mathbb{R}^n \\ (1 - \alpha)f(y) &\geq (1 - \alpha)f(\bar{x}) + \langle (1 - \alpha)d^2, y - \bar{x} \rangle + (1 - \alpha) \theta_2(\|y - \bar{x}\|) \quad \forall y \in \mathbb{R}^n. \end{aligned}$$

Sumando ambas desigualdades tenemos que

$$f(y) \geq f(\bar{x}) + \langle \alpha d^1 + (1 - \alpha)d^2, y - \bar{x} \rangle + \theta(\|y - \bar{x}\|) \quad \forall y \in \mathbb{R}^n, \quad (3.1)$$

donde $\theta(\|y - \bar{x}\|) = \alpha \theta_1(\|y - \bar{x}\|) + (1 - \alpha) \theta_2(\|y - \bar{x}\|)$. Dado que

$$\lim_{\substack{y \rightarrow \bar{x} \\ y \neq \bar{x}}} \frac{\theta(\|y - \bar{x}\|)}{\|y - \bar{x}\|} = 0,$$

entonces en (3.1) tenemos que $\alpha d^1 + (1 - \alpha)d^2 \in \widehat{\partial}f(\bar{x})$ para todo $\alpha \in [0, 1]$.

(b) Probaremos que $\widehat{\partial}f(\bar{x})$ es cerrado. Sea $\{d^k\}$ una sucesión en $\widehat{\partial}f(\bar{x})$ tal que $d^k \rightarrow d$. Como $d^k \in \widehat{\partial}f(\bar{x})$, entonces $f(y) \geq f(\bar{x}) + \langle d^k, y - \bar{x} \rangle + \theta(\|y - \bar{x}\|)$. Tomando límite cuando k va para el infinito, y usando la continuidad del producto interno $\langle \cdot, \cdot \rangle$ tenemos que $f(y) \geq f(\bar{x}) + \langle d, y - \bar{x} \rangle + \theta(\|y - \bar{x}\|)$, para todo $y \in \mathbb{R}^n$. Así que $d \in \widehat{\partial}f(\bar{x})$.

(c) Probaremos que $\partial^{Lim}f(\bar{x})$ es cerrado. Sea $\{d^k\}$ una sucesión en $\partial^{Lim}f(\bar{x})$ tal que $d^k \rightarrow d$. Como $d^k \in \partial^{Lim}f(\bar{x})$, entonces existe una sucesión $\{x^{l_k}\}$ tal que $x^{l_k} \xrightarrow{f} \bar{x}$; también existe $d^{l_k} \in \widehat{\partial}f(x^{l_k})$ tal que $d^{l_k} \rightarrow d^k$. Dado que $d^{l_k} \in \widehat{\partial}f(x^{l_k})$, entonces

$$f(y) \geq f(x^{l_k}) + \langle d^{l_k}, y - x^{l_k} \rangle + \theta(\|y - x^{l_k}\|) \quad \forall y \in \mathbb{R}^n,$$

haciendo que $l_k \rightarrow +\infty$ tenemos que

$$f(y) \geq f(\bar{x}) + \langle d^k, y - \bar{x} \rangle + \theta(\|y - \bar{x}\|) \quad \forall y \in \mathbb{R}^n,$$

lo cual implica que $d^k \in \widehat{\partial}f(\bar{x})$. Haciendo que $\{x^k\} = \bar{x}$, tenemos que $x^k \xrightarrow{f} \bar{x}$, y teniendo en cuenta que $d^k \in \widehat{\partial}f(x^k)$ converge a d , se tiene que $d \in \partial^{Lim}f(\bar{x})$.

(d) Finalmente probaremos que $\widehat{\partial}f(\bar{x}) \subset \partial^{Lim}f(\bar{x})$. Sea $d \in \widehat{\partial}f(\bar{x})$, entonces $f(y) \geq f(\bar{x}) + \langle d, y - \bar{x} \rangle + \theta(\|y - \bar{x}\|)$ para todo $y \in \mathbb{R}^n$. Haciendo que $\bar{x} = \{x^k\}$ y $d = \{d^k\}$, tenemos $x^k \xrightarrow{f} \bar{x}$ y $d^k \rightarrow d$, inmediatamente se tiene que $d \in \partial^{Lim}f(\bar{x})$. \square

Proposición 3.1.2. (Regla generalizada de Fermat). Si una función propia $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ tiene un mínimo local en $\bar{x} \in \text{dom}(f)$, entonces $0 \in \widehat{\partial}f(\bar{x})$.

Demostración. Dado que f tiene un mínimo local en \bar{x} , entonces existe $\epsilon_{\bar{x}} > 0$ tal que

$$f(\bar{x}) \leq f(y) \quad \forall y \in B(\bar{x}, \epsilon_{\bar{x}}),$$

donde $B(\bar{x}, \epsilon_{\bar{x}}) = \{y \in \mathbb{R}^n : \|y - \bar{x}\| < \epsilon_{\bar{x}}\}$. Definiendo $\theta(\|y - \bar{x}\|) = 0$, en la expresión anterior tenemos que

$$f(y) \geq f(\bar{x}) + \langle 0, y - \bar{x} \rangle + \theta(\|y - \bar{x}\|) \quad \forall y \in B(\bar{x}, \epsilon_{\bar{x}}).$$

Dado que

$$\lim_{\substack{y \rightarrow \bar{x} \\ y \neq \bar{x}}} \frac{\theta(\|y - \bar{x}\|)}{\|y - \bar{x}\|} = 0,$$

entonces por la Definición 3.1.1 (1) podemos concluir que $0 \in \widehat{\partial}f(\bar{x})$. \square

Las proposiciones que a continuación presentamos, no serán demostradas, dado que la realización de las mismas implican muchísimos conceptos, los cuáles escapan de los fines de la presente tesis.

Proposición 3.1.3. Sea $C \subset \mathbb{R}^n$, $C \neq \emptyset$, un conjunto convexo cerrado y $x \in C$, entonces

$$\partial^{Lim}(\delta_C)(x) = N_C(x) \quad (3.2)$$

Demostración. La demostración es una implicación directa del Ejercicio 4.5 (página 61) y la Proposición 2.9 (página 30) de Clarke (2013). \square

Proposición 3.1.4. Sean $f, g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ funciones propias, tales que f es localmente Lipschitz en $\bar{x} \in \text{dom}(f) \cap \text{dom}(g)$, y g es semicontinua inferior en dicho punto, entonces

$$\partial^{Lim}(f + g)(\bar{x}) \subset \partial^{Lim}f(\bar{x}) + \partial^{Lim}g(\bar{x}).$$

Demostración. Véase Mordukhovich (2006), Teorema 2.33. \square

3.1.2. Subdiferencial de Clarke

Definición 3.1.2. Sean $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ una función propia y localmente Lipschitz en $x \in \text{dom}(f)$, y $d \in \mathbb{R}^n$.

1. **La derivada direccional de Clarke** de f en x en la dirección d , denotada por $f^0(x, d)$, es definida como:

$$f^0(x, d) = \lim_{\substack{y \rightarrow x \\ t \downarrow 0}} \sup \frac{f(y + td) - f(y)}{t}.$$

2. **El subdiferencial de Clarke** de f en x , denotado por $\partial^0 f(x)$ está definido por

$$\partial^0 f(x) = \{w \in \mathbb{R}^n : \langle w, d \rangle \leq f^0(x, d), \quad \forall d \in \mathbb{R}^n\}.$$

Observación 3.1.1. De la definición anterior se sigue directamente que para todo $x \in \text{dom}(f)$ se tiene que $\widehat{\partial}f(x) \subset \partial^{Lim}f(x) \subset \partial^0 f(x)$ Bolte et al. (ver 2007, inclusión (7)).

Ejemplo 3.1.1. Clarke (Ejemplo 2.1.3 1990, página 28). Calcular el subdiferencial de Clarke de la función $f : \mathbb{R} \rightarrow \mathbb{R}$ definida por $f(x) = |x|$.

Como f es localmente Lipschitz³, procedemos a calcular $\partial^0 f(x)$, para lo cual consideramos tres casos:

- Si $x > 0$, entonces

$$f^0(x, d) = \lim_{\substack{y \rightarrow x \\ t \downarrow 0}} \sup \frac{|y + td| - |y|}{t} = \begin{cases} d & \text{si } d \geq 0 \\ d & \text{si } d < 0 \wedge y + td \geq 0, \end{cases}$$

luego $f^0(0, d) = d$. Así que $\partial^0 f(x) = \{w \in \mathbb{R} : w.d \leq d, \forall d \in \mathbb{R}\}$, de esta manera $\partial^0 f(x)$ se reduce a $\{1\}$.

- Si $x < 0$, entonces

$$f^0(x, d) = \lim_{\substack{y \rightarrow x \\ t \downarrow 0}} \sup \frac{|y + td| - |y|}{t} = \begin{cases} -d & \text{si } d \geq 0 \wedge y + td < 0 \\ -d & \text{si } d < 0, \end{cases}$$

de modo que $\partial^0 f(x) = \{w \in \mathbb{R} : w.d \leq -d, \forall d \in \mathbb{R}\}$, de donde se deduce que $\partial^0 f(x) = \{-1\}$.

- Finalmente si $x = 0$, se tiene que

$$f^0(x, d) = \lim_{\substack{y \rightarrow x \\ t \downarrow 0}} \sup \frac{|y + td| - |y|}{t} = \begin{cases} d & \text{si } d \geq 0 \\ -d & \text{si } d < 0, \end{cases}$$

luego $f^0(0, d) = |d|$. Así que $\partial^0 f(d) = \{w \in \mathbb{R} : w.d \leq |d|, \forall d \in \mathbb{R}\}$, por lo tanto $\partial^0 f(x) = [-1, 1]$.

Ejemplo 3.1.2. La función $f(x) = x^2 \sin(\frac{1}{x})$ es diferenciable, excepto en el punto cero. Hallaremos el subdiferencial de Clarke en este punto. Su gráfica es:

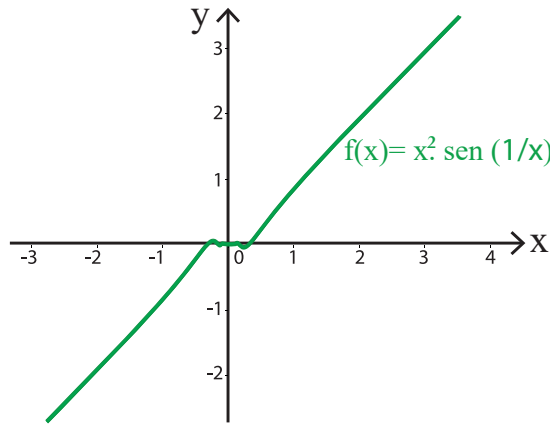


Figura 3.1: Gráfica de la función $f(x) = x^2 \sin(1/x)$.

³Esto inmediatamente se sigue de la desigualdad triangular: $|f(y) - f(x)| = ||y| - |x|| \leq |y - x|$.

Usando la definición de derivada generalizada tenemos:

$$f^0(x, d) = \lim_{\substack{y \rightarrow x \\ t \downarrow 0}} \sup \frac{(y + td)^2 \sin(\frac{1}{y+td}) - y^2 \sin(\frac{1}{y})}{t}$$

Tomando $x = 0$, y haciendo uso de la Definición 3.1.2(2) tenemos:

$$\begin{aligned} \partial^0 f(0) &= \{w \in \mathbb{R} : w.d \leq f^0(0, d), \quad \forall d \in \mathbb{R}\} \\ &= \left\{ w \in \mathbb{R} : w.d \leq \lim_{\substack{y \rightarrow 0 \\ t \downarrow 0}} \sup \frac{f(y + td) - f(y)}{t} \right\} \\ &= \left\{ w \in \mathbb{R} : w.d \leq \lim_{\substack{y \rightarrow 0 \\ t \downarrow 0}} \sup \frac{(y + td)^2 \sin(\frac{1}{y+td}) - y^2 \sin(\frac{1}{y})}{t} \right\} \\ &= \left\{ w \in \mathbb{R} : w.d \leq \lim_{\substack{y \rightarrow 0 \\ t \downarrow 0}} \sup \frac{t^2 d^2 \cdot \sin(\frac{1}{t.d})}{t} \right\} \\ &= \left\{ w \in \mathbb{R} : w.d \leq \lim_{\substack{y \rightarrow 0 \\ t \downarrow 0}} \sup \frac{d \cdot \sin(\frac{1}{t.d})}{\frac{1}{t.d}} \right\} \\ &= \{w \in \mathbb{R} : w.d \leq d\} \end{aligned}$$

i) Si $d \geq 0$, entonces se tiene que $w \leq 1$;

ii) si $d < 0$, entonces se tiene que $w \geq -1$.

De (i) y (ii) se tiene que $w \in [-1, 1]$, luego $\partial^0 f(0) = [-1, 1]$.

Lema 3.1.1. Sean $f, g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ funciones propias y localmente Lipschitz en $x \in \text{dom}(f) \cap \text{dom}(g)$. Luego para todo $d \in \mathbb{R}^n$ se tiene que

i) $(f + g)^0(x, d) \leq f^0(x, d) + g^0(x, d)$;

ii) $(\lambda f)^0(x, d) = \lambda f^0(x, d)$, para todo $\lambda \geq 0$;

iii) $f^0(x, \lambda d) = \lambda f^0(x, d)$, para todo $\lambda \geq 0$.

Demostración. (i). Por la Definición 3.1.2(1) es claro que

$$\begin{aligned} (f + g)^0(x, d) &= \lim_{\substack{y \rightarrow x \\ t \downarrow 0}} \sup \frac{(f + g)(y + td) - (f + g)(y)}{t} \\ &= \inf_{\delta > 0} \sup_{y \in B(x, \delta)} \frac{(f + g)(y + td) - (f + g)(y)}{t} \\ &= \inf_{\delta > 0} \sup_{y \in B(x, \delta)} \left(\frac{f(y + td) - f(y)}{t} + \frac{g(y + td) - g(y)}{t} \right). \end{aligned}$$

Dado que $\sup(\phi_1 + \phi_2 + \dots)(\cdot) \leq \sup(\phi_1)(\cdot) + \sup(\phi_2)(\cdot) + \dots$, entonces en la expresión anterior tenemos que

$$\begin{aligned} (f + g)^0(x, d) &\leq \inf_{\delta > 0} \left(\sup_{y \in B(x, \delta)} \frac{f(y + td) - f(y)}{t} + \sup_{y \in B(x, \delta)} \frac{g(y + td) - g(y)}{t} \right) \\ &= \inf_{\delta > 0} \sup_{y \in B(x, \delta)} \frac{f(y + td) - f(y)}{t} + \inf_{\delta > 0} \sup_{y \in B(x, \delta)} \frac{g(y + td) - g(y)}{t}, \end{aligned}$$

de donde inmediatamente se sigue el resultado. La relación (ii) se sigue directamente de la Definición 3.1.2(1).

(iii). Si $\lambda = 0$, la igualdad se verifica trivialmente, pero si $\lambda > 0$, por la Definición 3.1.2(1) tenemos que:

$$\begin{aligned} \lambda f^0(x, \lambda d) &:= \lim_{\substack{y \rightarrow x \\ t \downarrow 0}} \sup \frac{\lambda f(y + td) - \lambda f(y)}{t} \\ &= \lim_{\substack{y \rightarrow x \\ \frac{t}{\lambda} \downarrow 0}} \sup \frac{f(y + \frac{t}{\lambda}(\lambda d)) - f(y)}{\frac{t}{\lambda}} = f^0(x, \lambda d), \end{aligned}$$

de donde se sigue el resultado. \square

Lema 3.1.2. Si $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ es una función propia y localmente Lipschitz en $x \in \text{dom}(f)$, y λ un escalar arbitrario, entonces:

$$\partial^0(\lambda f)(x) = \lambda \partial^0 f(x). \quad (3.3)$$

Demostración. Dado que f es localmente Lipschitz, entonces λf también es localmente Lipschitz. Probaremos que:

i) $\partial^0(\lambda f)(x) \subset \lambda \partial^0 f(x)$. Sea $w \in \partial^0(\lambda f)(x)$, entonces por la Definición 3.1.2(2) se tiene que

$$\langle w, d \rangle \leq (\lambda f)^0(x, d) \quad \forall d \in \mathbb{R}^n. \quad (3.4)$$

Teniendo en cuenta el signo de λ , analizamos los siguientes casos:

i.1) Si $\lambda \geq 0$, por el Lema 3.1.1 (ii) se tiene que $(\lambda f)^0 = \lambda f^0$, entonces en (3.4)

$$\langle w, d \rangle \leq \lambda f^0(x, d) \quad \forall d \in \mathbb{R}^n. \quad (3.5)$$

Si $\lambda = 0$, la igualdad (3.3) se verifica trivialmente, en cambio si $\lambda > 0$ tenemos que

$$\left\langle \frac{w}{\lambda}, d \right\rangle \leq f^0(x, d) \quad \forall d \in \mathbb{R}^n,$$

lo cual implica que $\frac{w}{\lambda} \in \partial^0 f(x)$, de donde se tiene que $w \in \lambda \partial^0 f(x)$.

- i.2) Para probar la otra situación en la cual $\lambda < 0$, es suficiente estudiar la fórmula para $\lambda = -1$. Sea $w \in \partial^0(-f)(x)$, entonces de forma análoga a (3.4) tenemos que

$$\langle w, d \rangle \leq (-f)^0(x, d) \quad \forall d \in \mathbb{R}^n. \quad (3.6)$$

Pero como

$$\begin{aligned} f^0(x, -d) &:= \limsup_{\substack{y \rightarrow x \\ t \downarrow 0}} \frac{f(y - td) - f(y)}{t} \quad (\text{hacemos } y - td = u) \\ &= \limsup_{\substack{u \rightarrow x \\ t \downarrow 0}} \frac{(-f)(u + td) - (-f)(u)}{t} = (-f)^0(x, d), \end{aligned}$$

entonces en (3.6)

$$\langle -w, -d \rangle = \langle w, d \rangle \leq f^0(x, -d) \quad \forall d \in \mathbb{R}^n$$

lo cual implica $-w \in \partial^0 f(x)$. Por lo tanto de (i.1) e (i.2) se tiene que $\partial^0(\lambda f)(x) \subset \lambda \partial^0 f(x)$ para todo $\lambda \in \mathbb{R}$.

ii) $\lambda \partial^0 f(x) \subset \partial^0(\lambda f)(x)$. Considerando un $w \in \partial^0 f(x)$, haciendo uso de la Definición 3.1.2(1), y siguiendo el camino inverso al realizado en el ítem (i), fácilmente se obtiene dicha inclusión. \square

Lema 3.1.3. Sean $f_i : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, $i = 1, \dots, p$, funciones propias y localmente Lipschitz en $x \in \cap_{i=1}^p \text{dom}(f_i)$, entonces:

$$\partial^0\left(\sum_{i=1}^p f_i\right)(x) \subset \sum_{i=1}^p \partial^0 f_i(x). \quad (3.7)$$

Demostración. La prueba se sigue inmediatamente a partir del Lema 3.1.1 (i). \square

Proposición 3.1.5. Si $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ es una función propia y localmente Lipschitz de constante $K > 0$ cerca de x , entonces f^0 es semicontinua superior (scs), es decir si $(x, d) \in \mathbb{R}^n \times \mathbb{R}^n$ y $\{(x^k, d^k)\}$ es una sucesión en $\mathbb{R}^n \times \mathbb{R}^n$ tal que $\lim_{k \rightarrow +\infty} (x^k, d^k) = (x, d)$, entonces

$$\limsup_{k \rightarrow +\infty} f^0(x^k, d^k) \leq f^0(x, d). \quad (3.8)$$

Demostración. Sean $\{x^k\}$ y $\{d^k\}$ sucesiones arbitrarias en $\mathbb{R}^n \times \mathbb{R}^n$ que convergen a x y d respectivamente. Por la definición de límite superior, tenemos que

$$\begin{aligned} f^0(x^k, d^k) &= \inf_{\delta > 0} \sup_{\|y - x^k\| + t < \delta} \left(\frac{f(y + td^k) - f(y)}{t} \right) \\ &\leq \sup_{\|y - x^k\| + t < \delta} \left(\frac{f(y + td^k) - f(y)}{t} \right), \quad \forall \delta > 0. \end{aligned}$$

Haciendo uso de la definición de supremo, y considerando que $\delta = \frac{1}{k}$, se tiene la existencia de las sucesiones $\{y^k\}$ y $\{t_k\} > 0$, tales que

$$\begin{aligned} f^0(x^k, d^k) &\leq \frac{f(y^k + t_k d^k) - f(y^k)}{t_k} + \frac{1}{k}, \\ \|y^k - x^k\| + t_k &< \frac{1}{k} \quad \forall k \in \mathbb{N}, \end{aligned}$$

de donde

$$\begin{aligned} f^0(x^k, d^k) - \frac{1}{k} &\leq \frac{f(y^k + t_k d^k) - f(y^k)}{t_k} = \\ &\quad \frac{f(y^k + t_k d) - f(y^k)}{t_k} + \frac{f(y^k + t_k d^k) - f(y^k + t_k d)}{t_k}. \end{aligned} \quad (3.9)$$

Nótese que el último término de la parte derecha de (3.9) está acotado por $K\|d^k - d\|$, lo cual se cumple gracias a que f es localmente Lipschitz en x con constante K (ver Definición 2.2.3), y por el hecho que $t_k > 0$; lo cual probamos a continuación:

$$\frac{1}{t_k} |f(y^k + t_k d^k) - f(y^k + t_k d)| \leq \frac{K}{t_k} \|(y^k + t_k d^k) - (y^k + t_k d)\| = \frac{K}{t_k} \|d^k - d\|,$$

Teniendo en cuenta la expresión anterior, en (3.9) se tiene que

$$f^0(x^k, d^k) - \frac{1}{k} \leq \frac{f(y^k + t_k d) - f(y^k)}{t_k} + K\|d^k - d\|.$$

Tomando límite superior cuando $k \rightarrow +\infty$, tenemos que

$$\limsup_{k \rightarrow +\infty} f^0(x^k, d^k) \leq f^0(x, d),$$

lo cual establece la semicontinuidad superior de f . □

Proposición 3.1.6. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ una función propia, cuasi-convexa y localmente Lipschitz en \mathbb{R}^n . Si $g \in \partial^0 f(x)$ y $f(y) < f(x)$, entonces $\langle g, y - x \rangle \leq 0$

Demostración. Sea $g \in \partial^0 f(x)$, entonces por la Definición 3.1.2(2)

$$\langle g, d \rangle \leq f^0(x, d) = \limsup_{\substack{y \rightarrow x \\ t \downarrow 0}} \frac{f(y + td) - f(y)}{t} \quad \forall d \in \mathbb{R}^n.$$

Haciendo $d = y - x$, y aplicando la definición de límite superior, en la expresión anterior tenemos que

$$\begin{aligned} \langle g, y - x \rangle \leq f^0(x, y - x) &= \inf_{\delta > 0} \sup_{y \in B(x, \delta), t \in (0, \delta)} \left(\frac{f(y + t(y - x)) - f(y)}{t} \right) \\ &\leq \sup_{y \in B(x, \delta), t \in (0, \delta)} \left(\frac{f(y + t(y - x)) - f(y)}{t} \right). \end{aligned} \quad (3.10)$$

Aplicando la definición de supremo a (3.10), y considerando que $\delta = \frac{\epsilon}{k}$ para todo $k \in \mathbb{N}$, se tiene la existencia de las sucesiones $\{t_k\} \subset (0, \frac{\epsilon}{k})$ y $\{y^k\} \subset B(x, \frac{\epsilon}{k})$ (entonces $y - y^k \in {}^4B(y - x, \epsilon)$), tales que $\|y_k - x\| + t_k < \frac{\epsilon}{k}$ para todo $k \in \mathbb{N}$, y además

$$\langle g, y - x \rangle \leq \sup_{\|y-x\|+t<\delta} \left(\frac{f(y+t(y-x)) - f(y)}{t} \right) \leq \frac{f(y^k + t_k(y - y^k)) - f(y^k)}{t_k} + \frac{\epsilon}{k},$$

de donde

$$\langle g, y - x \rangle \leq \frac{f((t_k)y + (1 - t_k)y^k) - f(y^k)}{t_k} + \frac{\epsilon}{k}.$$

Dado que f es cuasi-convexa y $f(y) < f(x)$, tenemos que

$$\langle g, y - x \rangle \leq \frac{f(x) - f(y^k)}{t_k} + \frac{\epsilon}{k},$$

tomando límite cuando $k \rightarrow +\infty$ obtenemos que: $\langle g, y - x \rangle \leq 0$. \square

Proposición 3.1.7. Si $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ es una función convexa, propia, y localmente Lipschitz de constante $K > 0$ en $x \in \mathbb{R}^n$, entonces $\partial^0 f(x)$ coincide con el subdiferencial en x (en el sentido del análisis convexo), y $f^0(x, d)$ coincide con la derivada direccional $f'(x, d)$ para cada d .

Demostración. i) Probaremos que $f^0(x, d) = f'(x, d)$.

i.1) De las definiciones 3.1.2(1) y 2.2.14 claramente se sigue que $f^0(x, d) \geq f'(x, d)$.

i.2) Para completar la demostración de (i), a continuación probaremos que $f^0(x, d) \leq f'(x, d)$. El primer término de la desigualdad anterior, también puede ser escrito como

$$f^0(x, d) := \lim_{\substack{y \rightarrow x \\ t \downarrow 0}} \sup \frac{f(y + td) - f(y)}{t} = \lim_{\epsilon \downarrow 0} \sup_{\|y-x\| < \epsilon\delta} \sup_{0 < t < \epsilon} \frac{f(y + td) - f(y)}{t},$$

donde δ es un número positivo fijo. Dado que f es convexa, por el Teorema 2.2.3 se sigue que la función

$$t \rightarrow \frac{f(y + td) - f(y)}{t}$$

es no decreciente, entonces

$$f^0(x, d) = \lim_{\epsilon \downarrow 0} \sup_{\|y-x\| < \epsilon\delta} \frac{f(y + \epsilon d) - f(y)}{\epsilon}. \quad (3.11)$$

⁴Como $y^k \in B(x, \frac{\epsilon}{k})$, entonces $\|y^k - x\| < \frac{\epsilon}{k}$, pero dado que $\|y^k - x\| = \|x - y^k\| = \|(\mathbf{y} - y^k) - (\mathbf{y} - x)\| < \frac{\epsilon}{k} \leq \epsilon$, entonces $y - y^k \in B(y - x, \epsilon)$.

Haciendo que $\phi(y) = \frac{f(y+\epsilon d) - f(y)}{\epsilon}$, por la condición de Lipschitz, para algún $y \in B(x, \epsilon\delta)$ se tiene que

$$\begin{aligned} |\phi(y) - \phi(x)| &\leq \left| \frac{f(y + \epsilon d) - f(x + \epsilon d)}{\epsilon} \right| + \left| \frac{f(x) - f(y)}{\epsilon} \right| \\ &\leq \frac{K}{\epsilon} \|y - x\| + \frac{K}{\epsilon} \|y - x\| = \frac{2K}{\epsilon} \|y - x\| \leq \frac{2K}{\epsilon} (\epsilon\delta) = 2K\delta. \end{aligned}$$

donde K es la constante de Lipschitz. Teniendo en cuenta propiedades básicas del valor absoluto, en la expresión anterior tenemos que

$$-2\delta K + \psi(x) \leq \psi(y) = \frac{f(y + \epsilon d) - f(y)}{\epsilon} \leq \psi(x) + 2\delta K.$$

Entonces en (3.11)

$$f^0(x, d) \leq \lim_{\epsilon \downarrow 0} \frac{f(x + \epsilon d) - f(x)}{\epsilon} + 2\delta K = f'(x, d) + 2\delta K.$$

Dado que δ es arbitrario, deducimos que $f^0(x, d) \leq f'(x, d)$. Entonces teniendo en cuenta (i.1) y (i.2) tenemos que $f^0(x, d) = f'(x, d)$.

ii) Para probar que $\partial^0 f(x) = \partial f(x)$, realizamos lo siguiente:

ii.1) $\partial^0 f(x) \subset \partial f(x)$. Sea $w \in \partial^0 f(x)$, entonces por la Definición 3.1.2(2) y la aplicación del ítem (i) se tiene que

$$\langle w, d \rangle \leq f^0(x, d) = f'(x, d) := \inf_{t>0} \frac{f(x + td) - f(x)}{t} \quad \forall d \in \mathbb{R}^n$$

Como lo anterior es válido para todo $d \in \mathbb{R}^n$, en particular también será válido para $d = y - x$. Entonces

$$\begin{aligned} \langle w, y - x \rangle &\leq \inf_{t>0} \frac{f(x + t(y - x)) - f(x)}{t} \quad \forall d \in \mathbb{R}^n \\ t\langle w, y - x \rangle &\leq f(x + t(y - x)) - f(x) \quad (\text{como } f \text{ es convexa}) \\ t\langle w, y - x \rangle &\leq t(f(y) - f(x)) \\ \langle w, y - x \rangle + f(x) &\leq f(y) \quad \forall y \in \mathbb{R}^n, \end{aligned}$$

de donde se tiene que $w \in \partial f(x)$ para todo $y \in \mathbb{R}^n$.

ii.2) $\partial f(x) \subset \partial^0 f(x)$. Tomando un $w \in \partial f(x)$, aplicando la Definición 2.2.15 (2), y siguiendo el camino inverso al realizado en la demostración anterior, fácilmente se obtiene que $w \in \partial^0 f(x)$ para todo $d \in \mathbb{R}^n$. \square

Teorema 3.1.1. Si $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ es una función localmente Lipschitz y continuamente diferenciable, entonces $f^0(x, d)$ coincide con la derivada direccional $f'(x, d)$ para cada d .

Demostración. Sean $\{y^i\}$ y $\{t_i\}$ sucesiones arbitrarias tales que $y^i \rightarrow x$ y $t_i \downarrow 0$, entonces

$$f^0(x, d) = \lim_{i \rightarrow +\infty} \frac{f(y^i + t_i d) - f(y^i)}{t_i}. \quad (3.12)$$

Dado que f es continuamente diferenciable, por el teorema del valor medio ⁵, se tiene la existencia de algún $z_i \in [y^i, y^i + t_i d]$, tal que

$$\frac{f(y^i + t_i d) - f(y^i)}{(y^i + t_i d) - y^i} = \frac{f(y^i + t_i d) - f(y^i)}{t_i \cdot d} = f'(z_i).$$

Teniendo en cuenta la expresión anterior, así como también la continuidad de f , en (3.12) tenemos que

$$\begin{aligned} f^0(x, d) &= \lim_{i \rightarrow +\infty} \frac{f(y^i + t_i d) - f(y^i)}{t_i} = \lim_{i \rightarrow +\infty} \langle f'(z_i), d \rangle \\ &= \langle f'(x), d \rangle = f'(x, d), \end{aligned}$$

de donde se sigue el resultado. \square

Lema 3.1.4. Sean $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ una función localmente Lipschitz, y $g : \mathbb{R}^n \rightarrow \mathbb{R}$ una función convexa, entonces:

$$(f + g)^0(x, d) = f^0(x, d) + g^0(x, d) \quad \forall x, d \in \mathbb{R}^n. \quad (3.13)$$

Además, si g es diferenciable se tiene que

$$\partial^0(f + g)(x) = \partial^0 f(x) + \nabla g(x) \quad \forall x \in \mathbb{R}^n. \quad (3.14)$$

Demostración. Usando la Definición 3.1.2(1) y realizando una simple manipulación algebraica, tenemos que

$$(f + g)^0(x, d) = \lim_{\substack{y \rightarrow x \\ t \downarrow 0}} \sup \left(\frac{f(y + td) - f(y)}{t} + \frac{g(y + td) - g(y)}{t} \right). \quad (3.15)$$

Por una propiedad básica de límite superior junto con la Definición 3.1.2(1), en la expresión anterior se tiene que

$$\begin{aligned} (f + g)^0(x, d) &\leq \lim_{\substack{y \rightarrow x \\ t \downarrow 0}} \sup \left(\frac{f(y + td) - f(y)}{t} \right) + \lim_{\substack{y \rightarrow x \\ t \downarrow 0}} \sup \left(\frac{g(y + td) - g(y)}{t} \right) = \\ &= f^0(x, d) + g^0(x, d), \end{aligned} \quad (3.16)$$

pero como g es convexa, entonces por la Proposición 3.1.7

$$(f + g)^0(x, d) \leq f^0(x, d) + g'(x, d) \quad \forall x, d \in \mathbb{R}^n. \quad (3.17)$$

⁵Si f es continua en $[a, b]$ y diferenciable en (a, b) , entonces existe un $c \in (a, b)$ tal que $f'(c) = \frac{f(b) - f(a)}{b - a}$.

Dado que $f^0(x, d) = ((f + g) + (-g))^0(x, d)$, aplicando la desigualdad anterior tenemos que

$$f^0(x, d) \leq (f + g)^0(x, d) + (-g)'(x, d),$$

así que

$$(f + g)^0(x, d) \geq f^0(x, d) + g'(x, d) \quad \forall x, d \in \mathbb{R}^n. \quad (3.18)$$

Combinando las desigualdades (3.17) y (3.18) se obtiene la ecuación (3.13).

En el caso que g sea diferenciable, en particular $g'(x, d) = (\nabla g(x), d)$. Por lo tanto la prueba de la desigualdad (3.14) es una consecuencia inmediata de (3.13) y de la Definición 3.1.2(1). \square

3.2. Direcciones de descenso

Ahora podemos introducir la definición de punto crítico Pareto-Clarke para funciones localmente Lipschitz en \mathbb{R}^n , definición que jugará un papel clave en el desarrollo de la presente tesis.

Definición 3.2.1. (Custódio et al., 2011, Definición 4.6). Sea $F : \mathbb{R}^n \rightarrow \mathbb{R}^p$ una función localmente Lipschitz en \mathbb{R}^n . Decimos que $\hat{x} \in \mathbb{R}^n$ es un **punto crítico Pareto-Clarke** de F si para todas las direcciones $d \in \mathbb{R}^n$, existe $i_0 = i_0(d) \in \{1, \dots, p\}$ tal que $F_{i_0}^0(\hat{x}, d) \geq 0$.

La definición anterior esencialmente nos dice que en \mathbb{R}^n no existe ninguna dirección de descenso para todas las funciones objetivo.

Observación 3.2.1. De la Definición 3.2.1 se sigue que si un punto x no es un **punto crítico Pareto-Clarke**, entonces existe una dirección $d \in \mathbb{R}^n$ que satisface

$$F_i^0(x, d) < 0 \quad \forall i \in \{1, \dots, p\},$$

lo cual implica que para cada $i \in \{1, \dots, p\}$, d es una dirección de descenso para cada función F_i , es decir existe un $\epsilon > 0$, tal que

$$F_i(x + td) < F_i(x) \quad \forall t \in (0, \epsilon] \text{ y } \forall i \in \{1, \dots, p\}.$$

Lo cual en función de las componentes de orden de la función multiobjetivo significa que

$$F(x + td) \prec F(x) \quad \forall t \in (0, \epsilon]. \quad (3.19)$$

Proposición 3.2.1. Sea \hat{x} un punto crítico Pareto-Clarke de una función localmente Lipschitz $F : \mathbb{R}^n \rightarrow \mathbb{R}^p$. Si F es \mathbb{R}_{\geq}^p -convexa, entonces \hat{x} es una solución débil de Pareto del problema (4.2).

Demostración. Como \hat{x} es un punto crítico Pareto-Clarke de F , entonces para todas las direcciones d , existe $i_0 = i_0(d) \in \{1, \dots, p\}$ tal que $F_{i_0}^0(\hat{x}, d) \geq 0$; y dado que F es \mathbb{R}_{\geq}^p -convexa, por la Proposición 3.1.7

$$F_{i_0}'(\hat{x}, d) \geq 0, \quad (3.20)$$

donde $F_{i_0}'(\hat{x}, d)$ denota la derivada direccional de la función convexa F_{i_0} en \hat{x} en la dirección d . Por otra parte supongamos que \hat{x} no es una solución débil de Pareto del problema (4.2), entonces existe $x^* \in \mathbb{R}^n$ tal que $F(x^*) \prec F(\hat{x})$ (ver Definición 4.3.3), es decir $F_i(x^*) < F_i(\hat{x})$ para todo $i \in \{1, \dots, p\}$; así que para cada i , existe $\alpha = \alpha(i) > 0$ tal que

$$F_i(x^*) + \alpha = F_i(\hat{x}).$$

Definamos $x_t = tx^* + (1-t)\hat{x}$, $t \in (0, 1)$. Haciendo uso de la igualdad anterior, y de la \mathbb{R}_{\geq}^p -convexidad de F , se tiene que

$$\begin{aligned} F_i(x_t) = F_i(tx^* + (1-t)\hat{x}) &\leq tF_i(x^*) + (1-t)F_i(\hat{x}) \\ F_i(tx^* + \hat{x} - t\hat{x}) &\leq t(F_i(\hat{x}) - \alpha) + F_i(\hat{x}) - tF_i(\hat{x}) = F_i(\hat{x}) - t\alpha, \end{aligned}$$

de donde

$$\frac{F_i(\hat{x} + t(x^* - \hat{x})) - F_i(\hat{x})}{t} \leq -\alpha < 0 \quad \forall t \in (0, 1). \quad (3.21)$$

Haciendo que $d = (x^* - \hat{x}) \in \mathbb{R}^n$, y tomando límite cuando $t \rightarrow 0^+$ se tiene que

$$\lim_{t \rightarrow 0^+} \left(\frac{F_i(\hat{x} + td) - F_i(\hat{x})}{t} \right) = F_i'(\hat{x}, d) < 0 \quad \forall t \in (0, 1) \text{ y } \forall i \in \{1, \dots, p\},$$

lo cual contradice la relación (3.20). Por lo tanto \hat{x} es una solución débil de Pareto del POM (4.2). \square

Capítulo 4

TEORÍA DE OPTIMIZACIÓN MULTI-OBJETIVO

En el presente capítulo nos limitaremos a la teoría básica necesaria para los fines de ésta tesis. Para una descripción más minuciosa referirse a Ehrgott (2005) o Sawaragi et al. (1985).

En optimización clásica generalmente se trabaja con problemas donde se requiere optimizar un solo objetivo, bien sea lineal o no lineal, sin embargo en las aplicaciones es común ver casos donde es necesario optimizar más de un objetivo, los cuáles por lo general se encuentran en conflicto entre sí, pues la mejora de uno de ellos, puede significar empeorar el desempeño de los otros. Los siguientes ejemplos ilustrarán más claramente lo mencionado.

Ejemplo 4.0.1. *(Problema de la compra de un automóvil - Ehrgott, 2005, Ejemplo 1.1). Supongamos que queremos comprar un automóvil, y que para ello tenemos identificado como opciones los siguientes modelos: VW Golf, Opel Astra, Ford Focus y Toyota Corolla. Nuestra decisión de compra se basa en considerar el precio, el consumo de petróleo y la potencia del automóvil. Supongamos que estamos interesados en un auto menor precio, menor consumo, y máxima potencia. Las características de cada modelo de automóvil se muestran en la siguiente tabla:*

Tabla 4.1: Características de los cuatro automóviles (caso discreto).

Criterios de compra	Alternativas			
	VW	Opel	Ford	Toyota
Precio (1000 Dólares)	16.2	14.9	14.0	14.9
Consumo (l /100km)	7.2	7.0	7.5	8.2
Poder (kW)	66.0	62.0	55.0	71.0

*Este es un problema de optimización multiobjetivo (POM), en el cuál intervienen tres funciones objetivo: F_1 (precio), F_2 (Consumo) y F_3 (potencia), y cuatro alternativas de decisión (VW, Opel, Ford y Toyota). Queremos seleccionar la “mejor” alternativa (**solución óptima**) que minimice simultáneamente F_1 , F_2 y $-F_3$ (maximizar F_3 es equivalente a minimizar $-F_3$).*

En la Tabla 4.4 podemos observar que el automóvil más poderoso es el que consume más petróleo, además éste no es el más económico, así que no podemos comprar un automóvil que sea poderoso, económico, y que a la vez consuma poco combustible (es decir, que minimice simultáneamente F_1 , F_2 y $-F_3$), de ésta forma sin ninguna información adicional, es imposible seleccionar una única solución. Sin embargo si en el problema consideramos una única función objetivo, tendremos un problema de optimización mono-objetivo, en el cuál podemos obtener directamente una única solución óptima. \square

A continuación definimos formalmente un problema de optimización multiobjetivo.

Definición 4.0.1. Un problema de optimización multiobjetivo (POM) puede formularse de manera general de la siguiente manera:

$$\min_{x \in X} F(x) = \min\{F(x) : x \in X\}, \quad (4.1)$$

donde:

- $F : \mathbb{R}^n \rightarrow \mathbb{R}^p$, con $F(x) = (F_1(x), \dots, F_p(x))$, $p \geq 2$, denota el vector de funciones objetivo, definido en el espacio variable o **espacio de decisión** \mathbb{R}^n , con valores en el **espacio objetivo** o espacio criterio \mathbb{R}^p ;
- $X \subseteq \mathbb{R}^n$ denota la región o conjunto factible, definido en el espacio de decisión;
- $x = (x_1, x_2, \dots, x_n)$ representa un vector columna de variables de decisión;
- $Y = F(X) \subset \mathbb{R}^p$ es el conjunto factible, definido en el espacio objetivo.

Observación 4.0.1. 1. Si en el problema (4.1) hacemos que $X = \mathbb{R}^n$, entonces estaremos ante su versión irrestricta, dada por:

$$\min_{x \in \mathbb{R}^n} F(x) = \min\{F(x) : x \in \mathbb{R}^n\}. \quad (4.2)$$

2. Si alguna de las funciones objetivo de (4.1) es de maximización, por ejemplo la función F_i , $i = 1, \dots, p$, entonces $(-F_i)$ será de minimización.

Ejemplo 4.0.2. Considerando el problema de optimización con dos objetivos y una variable de decisión:

$$\min_{x \in \mathbb{R}} F(x) = \min_{x \in \mathbb{R}} (e^x, (e^x - 3)^2 + 1), \quad (4.3)$$

podemos identificar que

- $F : \mathbb{R} \rightarrow \mathbb{R}^2$, denota el vector de funciones objetivo, definido por $F(x) = (e^x, (e^x - 3)^2 + 1) = (F_1(x), F_2(x))$;

- \mathbb{R} denota espacio de decisión ¹;
- $Y = F(\mathbb{R})$ denota el conjunto factible, definido en el espacio objetivo \mathbb{R}^2 .

Si consideramos individualmente cada una de las componentes de la función F , como se muestra en la Figura 4.1, fácilmente se puede observar que la función F_1 tiene ínfimo, pero no mínimo global(ver las explicaciones relacionadas a la Figura 2.6 - Página 25); y que la función F_2 tiene como minimizador al punto $(\ln(3), 1)$. Sin embargo, considerando ambas funciones a la vez (como así lo exige el problema), la pregunta natural que uno podría plantearse es: **¿cuál o cuáles son los valores que debe asumir la variable x , de tal manera que minimice ambas funciones a la vez?**. En el resto de este capítulo mostraremos una forma adecuada de dar respuesta a tal interrogante. \square

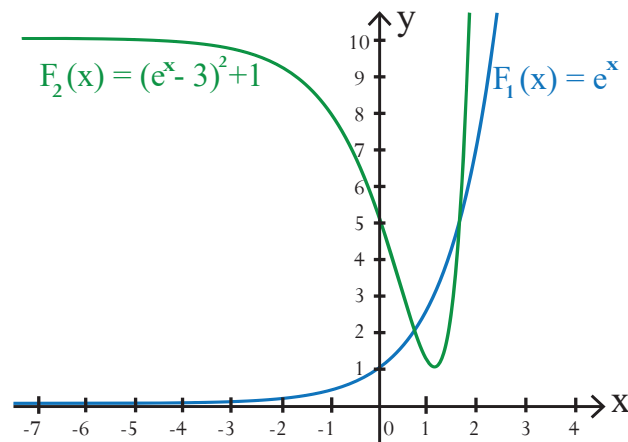


Figura 4.1: Gráfica de las funciones objetivo del Ejemplo 4.0.2.

4.1. Ordenes de preferencia

Puesto que en los problemas de optimización multiobjetivo (P_sOM), los objetivos suelen estar en conflicto, entonces raramente se puede esperar la existencia de una única solución que optimice de manera simultánea a todos estos objetivos; por lo tanto existirá un conjunto de soluciones, consideradas (sin información adicional) igualmente válidas para el problema. La elección de al menos una de estas soluciones, implica un proceso de negociación, en el cuál las actitudes de preferencia del tomador de decisiones desempeñan un papel esencial, pues especifican el significado de “optimalidad” o “deseabilidad”, muy a menudo es representada como una “**relación binaria**” en el espacio objetivo $Y \subset \mathbb{R}^p$.

Dotando de ciertas propiedades a una relación, se tendrá un “**orden de preferencia**”, el cual es fundamental en el contexto de optimización multiobjetivo,

¹Si el problema (4.3) lo hubiésemos planteado de la siguiente forma: $\min_{x \geq 0} F(x)$, se tendría que $X = \{x \in \mathbb{R} : x \geq 0\}$ representa el conjunto factible, definido en el espacio de decisión \mathbb{R} .

pues permite ordenar los elementos de un conjunto, en función de un orden previamente establecido.

En la presente sección introducimos las relaciones binarias, y presentamos algunas de sus propiedades, las cuáles nos van a permitir definir varias clases de órdenes, entre los que figuran los ordenes parcial y parcial estricto, considerados como base de las más importantes clases de relaciones que se dan en optimización multiobjetivo.

Definición 4.1.1. Sea Y cualquier conjunto en el espacio \mathbb{R}^P . Una **relación binaria** en Y es un subconjunto \mathcal{R} del producto cartesiano $Y \times Y$.

Escribiremos $y^1 \mathcal{R} y^2$ para indicar que $(y^1, y^2) \in \mathcal{R}$ o que y^1 está relacionado con y^2 , y para indicar que $(y^1, y^2) \notin \mathcal{R}$ escribiremos $y^1 \not\mathcal{R} y^2$.

Definición 4.1.2. Una relación binaria \mathcal{R} en un conjunto Y es llamada:

1. Reflexiva, si $(y, y) \in \mathcal{R}$ para todo $y \in Y$,
2. Simétrica, si $(y^1, y^2) \in \mathcal{R} \Rightarrow (y^2, y^1) \in \mathcal{R}$, para cualquiera $y^1, y^2 \in Y$,
3. Asimétrica, si $(y^1, y^2) \in \mathcal{R} \Rightarrow (y^2, y^1) \notin \mathcal{R}$, para cualquiera $y^1, y^2 \in Y$,
4. Antisimétrica, si $(y^1, y^2) \in \mathcal{R}$ y $(y^2, y^1) \in \mathcal{R} \Rightarrow y^1 = y^2$, para cualquiera $y^1, y^2 \in Y$,
5. Transitiva, si $(y^1, y^2) \in \mathcal{R}$ y $(y^2, y^3) \in \mathcal{R} \Rightarrow (y^1, y^3) \in \mathcal{R}$, para cualquiera $y^1, y^2, y^3 \in Y$, y
6. Conectada, conexa o completa, si $(y^1, y^2) \in \mathcal{R}$ o $(y^2, y^1) \in \mathcal{R}$, para cualquiera $y^1, y^2 \in Y$.

Ejemplo 4.1.1. Considerando el conjunto $A = \{1, 2, 3\}$, y la relación $\mathcal{R} = \{(1, 2); (2, 2); (3, 4); (4, 1)\} \subset A \times A$, fácilmente se puede verificar que \mathcal{R} no es simétrica, en vista de que $(1, 2) \in \mathcal{R}$, pero $(2, 1) \notin \mathcal{R}$. Además, \mathcal{R} no es asimétrica. \square

Definición 4.1.3. Una relación binaria \mathcal{R} sobre un conjunto Y se dice que es un **preorden** si es reflexiva y transitiva.

En el contexto de (pre) órdenes otras notaciones para la relación \mathcal{R} son convenientes. La relación binaria “ \preceq ”, que de aquí en adelante asumiremos, para referirnos indirectamente a la relación \mathcal{R} , significa “preferido a”. Es decir $y^1 \preceq y^2$, para $y^1, y^2 \in Y$ implica que y^1 precede o es anterior a y^2 , o que el resultado (valor objetivo) de y^1 es preferible al de y^2 . En el caso que \preceq sea un preorden, entonces el par (Y, \preceq) será llamado un **conjunto preordenado**.

Asociada al preorden \preceq , se tiene la relación de estricta preferencia “ \prec ” definida por:

$$y^1 \prec y^2 \Leftrightarrow y^1 \preceq y^2 \text{ y } y^2 \not\preceq y^1. \quad (4.4)$$

Definición 4.1.4. Una relación binaria \preceq en Y se dice que es

1. un **orden parcial**, si es un preorden antisimétrico,
2. un **orden total**, si se trata de un orden parcial conectado²,
3. un **orden parcial estricto**, si es asimétrica y transitiva.

Un conjunto Y junto con una relación de orden total (parcial) \preceq , se llama **conjunto totalmente (parcialmente) ordenado** con respecto a dicha relación, y se denota por (Y, \preceq) . A los conjuntos totalmente ordenados también se les conoce como cadenas. Por ejemplo los conjuntos $(\mathbb{Z}$ y $\mathbb{N})$ dotados de alguna de las relaciones $(\geq$ o $\leq)$, son conjuntos totalmente ordenados.

Ejemplo 4.1.2. La relación \leq definida en \mathbb{Z} , representa una relación de orden total.

En efecto, sean $a, b \in \mathbb{Z}$, entonces:

$$\begin{aligned} a \leq b &\Leftrightarrow b - a \geq 0, \text{ o lo que es lo mismo} \\ a \leq b &\Leftrightarrow b - a \in \mathbb{Z}_0^+, \text{ es decir} \\ a \leq b &\Leftrightarrow \exists k \in \mathbb{Z}_0^+ \text{ tal que } b - a = k. \end{aligned}$$

Para afirmar que ésta relación representa una relación de orden total, probaremos que:

i) \leq representa un orden parcial (ver Definición 4.1.4 (1)). Veamos si esta relación cumple con las condiciones exigidas para ser considerada un orden parcial:

i.1. Reflexividad. En efecto, sea a un elemento arbitrario de \mathbb{Z} , entonces:

$$a = a \implies a - a = 0; \quad 0 \in \mathbb{Z}_0^+ \implies a \leq a.$$

i.2. Antisimetría. En efecto, sean a y b dos enteros cualesquiera, entonces:

$$\begin{aligned} \left. \begin{aligned} a \leq b &\Leftrightarrow \exists k_1 \in \mathbb{Z}_0^+ : b - a = k_1, \text{ y} \\ b \leq a &\Leftrightarrow \exists k_2 \in \mathbb{Z}_0^+ : a - b = k_2 \end{aligned} \right\} &\implies k_1 = -k_2, \quad k_1, k_2 \in \mathbb{Z}_0^+ \\ &\implies k_1 = k_2 = 0. \end{aligned}$$

Por lo tanto $b - a = a - b$, es decir $a = b$. Consecuentemente la relación es antisimétrica.

²Sean y^1 y y^2 dos elementos de un conjunto Y , sobre el que se ha definido una relaciona de orden \preceq , decimos que y^1 y y^2 están conectados o son comparables si uno de ellos es anterior al otro (ver Definición 4.1.2- ítem 6).

i.3. **Transitividad.** En efecto, sean a, b y c tres números enteros cualesquiera, entonces

$$\left. \begin{array}{l} a \leq b \iff \exists k_1 \in \mathbb{Z}_0^+ : b - a = k_1 \text{ y} \\ b \leq c \iff \exists k_2 \in \mathbb{Z}_0^+ : c - b = k_2 \end{array} \right\} \implies c - a = k_1 + k_2 = k \in \mathbb{Z}_0^+ \\ \implies a \leq c.$$

Por lo tanto, la relación es transitiva.

Dado que la relación \leq es reflexiva, transitiva y antisimétrica, entonces dicha relación representa un orden parcial.

ii) \leq satisface la propiedad **conexa** (ver Definición 4.1.2 - ítem 6), propiedad que implica que todos los números enteros son comparables por la relación \leq .

Consideramos dos números enteros arbitrarios a y b , probaremos que $a \leq b$ o $b \leq a$. En efecto, como a y b han sido elegidos arbitrariamente, puede ocurrir que sean iguales o distintos. Pues bien,

$$\begin{aligned} a = b \vee b \neq a &\iff a = b \vee b - a \neq 0 \\ &\iff a = b \vee b - a \in \mathbb{Z} \setminus \{0\} \\ &\iff a = b \vee b - a \in \mathbb{Z}^- \cup \mathbb{Z}^+ \\ &\iff a = b \vee \begin{cases} \exists k \in \mathbb{Z}^- : b - a = k \\ \vee \\ \exists k \in \mathbb{Z}^+ : b - a = k \end{cases} \\ &\iff a = b \vee \begin{cases} a - b = -k, -k \in \mathbb{Z}^+ \\ \vee \\ b - a = k, k \in \mathbb{Z}^+ \end{cases} \\ &\iff a = b \vee \begin{cases} b < a \\ \vee \\ a < b \end{cases} \\ &\iff \begin{cases} a = b \vee b < a \\ \vee \\ a = b \vee a < b \end{cases} \\ &\iff \begin{cases} b \leq a \\ \vee \\ a \leq b \end{cases} \end{aligned}$$

Dado que la relación de orden \leq definida en \mathbb{Z} satisface (i) y (ii), entonces concluimos que dicha relación es de orden total. \square

Los ordenes de preferencia que usualmente asumiremos a lo largo de éste trabajo, serán al menos los órdenes considerados en la Definición 4.1.4 - ítems 1. y 3.,

los cuáles están definidos sobre el espacio Euclidiano \mathbb{R}^p , $p \geq 2$. Los nombres y notaciones que utilizaremos para referirnos a dichos órdenes serán los que se detallan en la Tabla 4.2. Cabe mencionar que dichas notaciones no son las únicas en la literatura de optimización multiobjetivo, por lo cual se recomienda verificar las definiciones, al consultar otras fuentes.

Tabla 4.2: Algunos órdenes en \mathbb{R}^p .

Notación	Definición	Nombre
$y^1 \leq y^2$	$y_k^1 \leq y_k^2; \quad k = 1, \dots, p$	Orden débil por componentes
$y^1 \leq y^2$	$y_k^1 \leq y_k^2; \quad k = 1, \dots, p; \quad y^1 \neq y^2$	Orden por componentes
$y^1 < y^2$	$y_k^1 < y_k^2; \quad k = 1, \dots, p$	Orden estricto por componentes

Tabla 4.3: Fuente: Ehrgott (2005), Tabla 1.2, página12.

Así, con las componentes de orden (débil y estricto), definimos los siguientes subconjuntos sobre el espacio Euclidiano \mathbb{R}^p , $p \geq 2$:

$$\mathbb{R}_{\leq}^p = \{y \in \mathbb{R}^p : y \geq 0\}, \text{ denota el ortante no negativo de } \mathbb{R}^p; \quad (4.5)$$

$$\mathbb{R}_{\geq}^p = \{y \in \mathbb{R}^p : y \geq 0\} = \mathbb{R}_{\leq}^p \setminus \{0\}; \quad (4.6)$$

$$\mathbb{R}_{>}^p = \{y \in \mathbb{R}^p : y > 0\} = \text{int}(\mathbb{R}_{\leq}^p), \text{ denota el ortante positivo de } \mathbb{R}^p. \quad (4.7)$$

Notar que para $p = 1$, $\mathbb{R}_{\geq}^p = \mathbb{R}_{>}^p$.

4.2. Conos y ordenes relacionados

En el caso de los problemas de optimización simples o mono-objetivo, el significado de minimizar una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$, sobre un conjunto $X \subseteq \mathbb{R}^n$, es claro: el punto óptimo $\hat{x} \in X$ es una solución a tal problema si $f(\hat{x}) \leq f(x)$ para todo $x \in X$. Como puede verse en este caso, el significado de mínimo está directamente relacionado con el orden canónico \leq en \mathbb{R} . Esto no ocurre en los problemas de optimización multiobjetivo, donde $F : \mathbb{R}^n \rightarrow \mathbb{R}^p$, $p \geq 2$, dado que no existe un orden canónico en \mathbb{R}^p ; por lo tanto el significado de mínimo para el problema (4.2) estará relacionado con un determinado orden en \mathbb{R}^p , el cuál previamente debe ser establecido.

A continuación presentamos la definición de cono en \mathbb{R}^p , y posteriormente mostramos que es posible definirlo en función de una relación de orden, o viceversa, lo cual nos va a permitir obtener uno de los conos más utilizados en el ámbito de optimización multiobjetivo, el cono R_{\leq}^p ³, el cuál será punteado, convexo y además contendrá al cero, si dicha relación (compatible con la adición y multiplicación escalar) representa un orden parcial (ver Definición 4.1.4 (1)).

³La importancia del cono R_{\leq}^p , definido en función de la relación de orden débil por componentes (ver Tabla 4.2), radica en que nos va a permitir derivar a una interpretación geométrica de las propiedades de los ordenes.

Definición 4.2.1. Un subconjunto $C \subseteq \mathbb{R}^p$ es llamado un cono, si $\alpha d \in C$ para todo $d \in C$ y para todo $\alpha \in \mathbb{R}$, $\alpha > 0$. Un cono C se dice que es:

1. **no trivial o propio**, si $C \neq \emptyset$ y $C \neq \mathbb{R}^p$;
2. **convexo**, si $\alpha d^1 + (1-\alpha)d^2 \in C$ para todo $d^1, d^2 \in C$ y para todo $0 \leq \alpha \leq 1$;
3. **cerrado**, si éste es cerrado como un subconjunto de \mathbb{R}^p ;
4. **punteado**, si para todo $d \in C$, $d \neq 0$, se cumple que $-d \notin C$, es decir $C \cap \{-C\} \subseteq \{0\}$.

Debido a la definición de cono, podemos ver que C es un **cono convexo**, si $d^1 + d^2 \in C$, para cualquiera $d^1, d^2 \in C$. En efecto dados $d^1, d^2 \in C$, por la definición de cono se tiene que $\alpha d^1 \in C$ y $(1-\alpha)d^2 \in C$ para todo $0 < \alpha < 1$, por ser C un cono. Por lo tanto, la cerradura de C bajo la adición es suficiente para la convexidad.

A lo largo de éste trabajo, solamente usaremos conos no triviales.

Ejemplo 4.2.1. La Figura 4.2 (a) muestra al cono $C = \mathbb{R}_{\geq}^2 = \{d \in \mathbb{R}_{\geq}^2 : d_k \geq 0; k = 1, 2\}$, el cual es un cono de elementos no negativos, determinado por el orden débil por componentes (ver Tabla 4.2). La Figura 4.2 (b) muestra un pequeño cono $C \subset \mathbb{R}_{\geq}^2$. \square

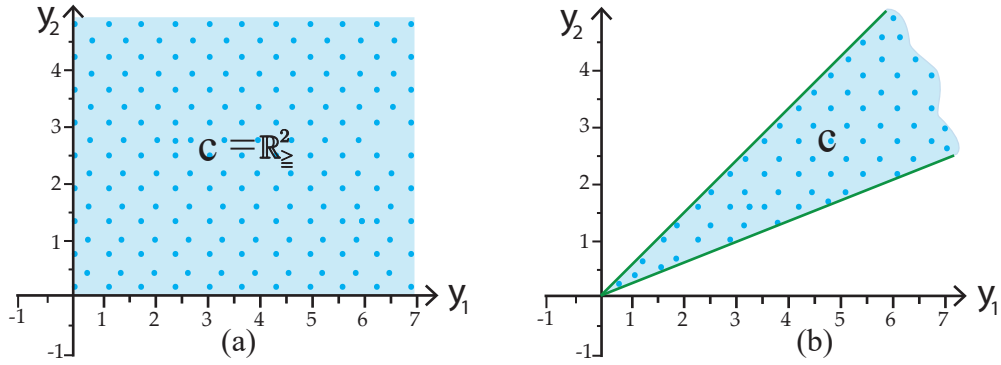


Figura 4.2: Ilustración de dos conos.

Definición 4.2.2. Sea \mathcal{R} una relación binaria sobre \mathbb{R}^p , definimos:

$$C_{\mathcal{R}} = \{y^2 - y^1 : y^1 \mathcal{R} y^2\}, \quad (4.8)$$

como el conjunto de elementos no negativos de \mathbb{R}^p de acuerdo con \mathcal{R} .

Ejemplo 4.2.2. Considerando el orden débil por componentes sobre \mathbb{R}^p (ver Tabla 4.2), dado por

$$\begin{aligned} y^1 \leq y^2 &\Leftrightarrow y_k^1 \leq y_k^2 \quad \forall k = 1, \dots, p \\ &\Leftrightarrow y_k^2 - y_k^1 = d_k \geq 0 \quad \forall k = 1, \dots, p, \end{aligned}$$

podemos definir el cono $C_{\leq} = \{d \in \mathbb{R}^p : d_k \geq 0, k = 1, \dots, p\} = \mathbb{R}_{\leq}^p$ (ver ecuación (4.5)). \square

Definición 4.2.3. Decimos que la relación binaria \mathcal{R} es

1. compatible con la multiplicación escalar, si para cualesquiera $(y^1, y^2) \in \mathcal{R}$ y para todo $\alpha \in \mathbb{R}_{>}$, se tiene que $(\alpha y^1, \alpha y^2) \in \mathcal{R}$,
2. compatible con la adición si para cualesquiera $y^1 \mathcal{R} y^2$ y para todo $z \in \mathbb{R}^p$, se tiene que $(y^1 + z) \mathcal{R} (y^2 + z)$.

Proposición 4.2.1. Si \mathcal{R} es compatible con la multiplicación escalar, entonces $C_{\mathcal{R}}$ definido en (4.8) es un cono.

Demostración. Para demostrar que $C_{\mathcal{R}}$ es un cono, tomando un elemento $d \in C_{\mathcal{R}}$, probaremos que $\alpha d \in C_{\mathcal{R}}$. En efecto, sea $d \in C_{\mathcal{R}}$ tal que $d = y^2 - y^1$, para cualesquiera $y^1, y^2 \in \mathbb{R}^p$, con $y^1 \mathcal{R} y^2$. Luego como \mathcal{R} es compatible con la multiplicación escalar, entonces $\alpha y^1 \mathcal{R} \alpha y^2$ para todo $\alpha > 0$. Por lo tanto $\alpha d = (\alpha y^2 - \alpha y^1) \in C_{\mathcal{R}}$ (pues $\alpha y^1 \mathcal{R} \alpha y^2$) para todo $\alpha > 0$. \square

Considerando la igualdad (4.8), es interesante ver que si fijamos el vector y^1 , entonces tendremos el conjunto:

$$C_{\mathcal{R}}(y^1) = \{y^2 - y^1 : y^1 \mathcal{R} y^2\}, \quad (4.9)$$

con el cuál definimos a $y^1 + C_{\mathcal{R}}(y^1)$ como el conjunto de elementos de \mathbb{R}^p dominados por y^1 (también conocida como la **región de dominancia de y^1**), cuyo concepto y propiedades serán presentados más adelante, pues son de suma importancia en la optimización multiobjetivo.

Una pregunta natural que uno podría plantearse respecto a la ecuación (4.9) es: ¿bajo qué condiciones $C_{\mathcal{R}}(y)$ es el mismo, para cualquier $y \in Y$? Con el fin de responder a tal interrogante, es necesario otra hipótesis sobre la relación de orden \mathcal{R} : la compatibilidad de \mathcal{R} con la adición.

Lema 4.2.1. Si \mathcal{R} es compatible con la adición y $d \in C_{\mathcal{R}}$, entonces $0 \mathcal{R} d$.

Demostración. Si $d \in C_{\mathcal{R}}$, entonces existen $y^1, y^2 \in \mathbb{R}^p$ con $y^1 \mathcal{R} y^2$ tal que $y^2 - y^1 = d$; haciendo $z = -y^1$ y usando la compatibilidad de \mathcal{R} con la adición (ver Definición 4.2.3 (2)), tenemos $(y^1 + z) \mathcal{R} (y^2 + z)$, o $0 \mathcal{R} d$. \square

Lo que nos dice este lema, es que si \mathcal{R} es compatible con la adición, entonces los conjuntos $C_{\mathcal{R}}(y)$, $y \in \mathbb{R}^p$, no dependen de y . A lo largo de esta tesis nos ocuparemos principalmente de éste caso.

Teorema 4.2.1. Sea \mathcal{R} una relación binaria sobre \mathbb{R}^p compatible con la adición y multiplicación escalar. Se cumplen las siguientes propiedades:

1. $0 \in C_{\mathcal{R}}$, si y sólo si \mathcal{R} es reflexiva.
2. $C_{\mathcal{R}}$ es punteado, si y sólo si \mathcal{R} es antisimétrica.
3. $C_{\mathcal{R}}$ es convexo, si y sólo si \mathcal{R} es transitiva.

Demostración. $(1 \leftarrow)$ Sean \mathcal{R} reflexiva y $y \in \mathbb{R}^p$, entonces $y\mathcal{R}y$, lo cual implica que $y - y = 0 \in C_{\mathcal{R}}$.

$(1 \rightarrow)$ Sea $0 \in C_{\mathcal{R}}$, entonces existe algún $y \in \mathbb{R}^p$ tal que $y\mathcal{R}y$; también existe $y' \in \mathbb{R}^p$ tal que $y' = y + z$, para algún $z \in \mathbb{R}^p$. Dado que \mathcal{R} es compatible con la adición, y del hecho que $y\mathcal{R}y$, entonces $(y + z)\mathcal{R}(y + z)$, o $y'\mathcal{R}y'$.

$(2 \leftarrow)$ Sean \mathcal{R} antisimétrica y $d \in C_{\mathcal{R}}$ tal que además $-d \in C_{\mathcal{R}}$, luego existen $y^1, y^2 \in \mathbb{R}^p$ tales que $y^1\mathcal{R}y^2$, donde $y^2 - y^1 = d$; también existen $y^3, y^4 \in \mathbb{R}^p$ tales que $y^3\mathcal{R}y^4$ y $y^4 - y^3 = -d$. Así tenemos que $y^2 - y^1 = y^3 - y^4$. También debe existir un $y \in \mathbb{R}^p$ tal que $y^3 + y = y^2$ y $y^4 + y = y^1$; luego como \mathbb{R}^p es compatible con la adición, entonces $(y^3 + y)\mathcal{R}(y^4 + y)$, o $y^2\mathcal{R}y^1$, luego por antisimetría de \mathcal{R} (utilizando las relaciones $y^1\mathcal{R}y^2$ y $y^2\mathcal{R}y^1$) tenemos que $y^1 = y^2$. Por lo tanto $d = 0$, es decir $C_{\mathcal{R}}$ es punteado.

$(2 \rightarrow)$ Sean $y^1, y^2 \in \mathbb{R}^p$ tales que $y^1\mathcal{R}y^2$ y $y^2\mathcal{R}y^1$, entonces $d = (y^2 - y^1) \in C_{\mathcal{R}}$ y $-d = (y^1 - y^2) \in C_{\mathcal{R}}$. Como $C_{\mathcal{R}}$ es punteado, entonces $\{d, -d\} \in C_{\mathcal{R}}$, implica que $d = 0$ y que $y^1 = y^2$. Por lo tanto \mathcal{R} es antisimétrica.

$(3 \leftarrow)$ Sean \mathcal{R} transitiva y $d^1, d^2 \in C_{\mathcal{R}}$; para probar que $C_{\mathcal{R}}$ es convexo, solamente necesitamos probar que $d^1 + d^2 \in C_{\mathcal{R}}$. Por el Lema 4.2.1 conocemos que $0\mathcal{R}d^1$ y $0\mathcal{R}d^2$; además, como \mathcal{R} es compatible con la adición, entonces $d^1\mathcal{R}(d^1 + d^2)$; luego por la transitividad de \mathcal{R} tenemos que $0\mathcal{R}(d^1 + d^2)$, lo cual implica que $(d^1 + d^2) \in C_{\mathcal{R}}$. Por lo tanto $C_{\mathcal{R}}$ es convexo.

$(3 \rightarrow)$ Sean $C_{\mathcal{R}}$ convexo y $y^1, y^2, y^3 \in \mathbb{R}^p$ tales que $y^1\mathcal{R}y^2$ y $y^2\mathcal{R}y^3$; probaremos que $y^1\mathcal{R}y^3$. Haciendo $d^1 = (y^2 - y^1) \in C_{\mathcal{R}}$ y $d^2 = (y^3 - y^2) \in C_{\mathcal{R}}$, y teniendo en cuenta la convexidad de $C_{\mathcal{R}}$ tenemos que $d^1 + d^2 = (y^3 - y^1) \in C_{\mathcal{R}}$; aplicando el Lema 4.2.1 a la expresión anterior tenemos que $0\mathcal{R}(y^3 - y^1)$, pues $y^3 - y^1 - 0 \in C_{\mathcal{R}}$; además, como \mathcal{R} es compatible con la adición, entonces $y^1\mathcal{R}y^3$. \square

Ejemplo 4.2.3. El orden débil por componentes \leq (ver Tabla 4.2) además de ser compatible con la adición y multiplicación escalar, determina un cono $C_{\leq} = \mathbb{R}_{\geq}^p$ punteado y convexo que contiene al cero. \square

En los puntos anteriores, hemos definido un cono C en función de una relación \mathcal{R} , denotándolo por $C_{\mathcal{R}}$. También es posible definir una relación de orden en función de un cono. Sea C un cono, definimos \mathcal{R}_C por:

$$y^1\mathcal{R}_C y^2 \iff y^2 - y^1 \in C. \quad (4.10)$$

Ejemplo 4.2.4. (orden de Pareto) Teniendo en cuenta la relación (4.10), considerando que $C = \mathbb{R}_{\geq}^p \setminus \{0\} = \mathbb{R}_{\geq}^p$, definimos una nueva relación binaria “ \leq ” en \mathbb{R}^p ($p \geq 2$) por:

$$\begin{aligned}
y^1 \leq y^2 &\iff y^2 - y^1 \in \mathbb{R}_{\geq}^p \setminus \{0\} = \mathbb{R}_{\geq}^p \\
&\iff y^2 - y^1 \in \mathbb{R}_{\geq}^p, y_1 \neq y_2 \\
&\iff y^2 - y^1 \geq 0, y_1 \neq y_2 \quad (\geq \text{denota el orden débil por componentes}) \\
&\iff y^1 \leq y^2, y_1 \neq y_2 \\
&\iff y_k^1 \leq y_k^2 \quad \forall k = 1, \dots, p, \text{ con } y_j^1 < y_j^2 \text{ para algún } j = 1, \dots, p. \quad (4.11)
\end{aligned}$$

Es claro que la relación binaria definida en (4.11) resulta ser un orden parcial estricto (ver Definición 4.1.4 (3)). Este orden “ \leq ” respecto del cono $C = \mathbb{R}_{\geq}^p = \mathbb{R}_{\geq}^p \setminus \{0\}$ se conoce como el orden de Pareto, y el concepto de optimización multiobjetivo respecto de este orden, se conoce como optimalidad de Pareto. La figura 4.3 ilustra el orden de Pareto en \mathbb{R}^2 . \square

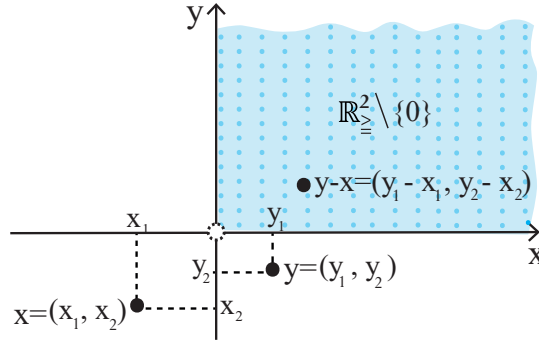


Figura 4.3: Ilustración del orden de Pareto “ \leq ” en \mathbb{R}^2 . Fuente Tablada (2011) - Figura 2.4.

Proposición 4.2.2. Sea C un cono, luego \mathcal{R}_C definido como en (4.10) es compatible con la adición y multiplicación escalar en \mathbb{R}^p .

Demostración. Sean $y^1, y^2 \in \mathbb{R}^p$ tales que $y^1 \mathcal{R}_C y^2$, luego $y^2 - y^1 = d \in C$.

- Dado que $((y^2 + z) - (y^1 + z)) \in C$, entonces $(y^1 + z) \mathcal{R}_C (y^2 + z)$ para todo $z \in \mathbb{R}^p$. Por lo tanto \mathcal{R}_C es compatible con la adición.
- Por otro lado, dado que C es un cono, entonces $\alpha d = (\alpha y^2 - \alpha y^1) \in C$, lo cual por la relación (4.10) implica que $(\alpha y^1) \mathcal{R}_C (\alpha y^2)$ para todo $\alpha > 0$. Por lo tanto \mathcal{R}_C también es compatible con la multiplicación escalar en \mathbb{R}^p . \square

Teorema 4.2.2. Sean C un cono y \mathcal{R}_C definido como en (4.10), entonces:

- \mathcal{R}_C es reflexiva si y solo si $0 \in C$.
- \mathcal{R}_C es antisimétrica si y solo si C es punteado.
- \mathcal{R}_C es transitiva si y solo si C es convexo.

Demostración. $(1 \leftarrow)$. Sean $0 \in C$ y $y \in \mathbb{R}^p$, entonces $y - y \in C$, lo cual implica que $y \mathcal{R}_C y$ para todo $y \in \mathbb{R}^p$.

$(1 \rightarrow)$. Sea \mathcal{R}_C reflexiva, entonces $y \mathcal{R}_C y$ para todo $y \in \mathbb{R}_{\geq}^p$, lo cual implica que $y - y = 0 \in C$.

$(2 \leftarrow)$. Sean $y^1, y^2 \in \mathbb{R}^p$ tales que $y^1 \mathcal{R}_C y^2$ y $y^2 \mathcal{R}_C y^1$ (asumiendo que Y es punteado, probaremos que $y^1 = y^2$). Consideremos $d = (y^2 - y^1) \in C$ y $-d = (y^1 - y^2) \in C$. Puesto que C es punteado, $d = 0$. Por lo tanto $y^1 = y^2$.

$(2 \rightarrow)$. Sean $d \in C$ y $-d \in C$, entonces $0 \mathcal{R}_C d$ y $0 \mathcal{R}_C -d$ (debemos probar que $d = 0$). Dado que \mathcal{R}_C es compatible con la adición, entonces $d \mathcal{R}_C 0$ (esto se da teniendo en cuenta la relación $0 \mathcal{R}_C -d$ y el hecho que $d \in C$); además, dado que \mathcal{R}_C es antisimétrica, entonces $d = 0$ (pues $0 \mathcal{R}_C d$ y $d \mathcal{R}_C 0$).

$(3 \leftarrow)$. Sean $y^1, y^2, y^3 \in \mathbb{R}^p$ tales que $y^1 \mathcal{R}_C y^2$ y $y^2 \mathcal{R}_C y^3$, entonces $d^1 = (y^2 - y^1) \in C$ y $d^2 = (y^3 - y^2) \in C$. Dado que C es convexo, entonces $d^1 + d^2 = (y^3 - y^1) \in C$, lo cual implica que $y^1 \mathcal{R}_C y^3$.

$(3 \rightarrow)$. Sean $d^1, d^2 \in C$, entonces por la relación (4.10) se tiene que $0 \mathcal{R}_C d^1$ y $0 \mathcal{R}_C d^2$. Dado que por \mathcal{R}_C es compatible con la adición (ver Proposición 4.2.2), entonces $d^1 \mathcal{R}_C (d^1 + d^2)$. Aplicando la transitividad de \mathcal{R}_C se tiene que $0 \mathcal{R}_C (d^1 + d^2)$ y $(d^1 + d^2 - 0) \in C$. \square

Con los Teoremas 4.2.1 y 4.2.2 se ha mostrado la equivalencia entre algunos órdenes parciales⁴ y los conos punteados y convexos que contienen al cero. Dado que los ordenes (parciales) pueden ser usados para definir “minimización”, éstos resultados hacen que sea posible analizar geométricamente problemas de optimización multiobjetivo (PsOM).

4.3. Eficiencia, débil eficiencia y no dominancia de soluciones

Como ya se mencionó, en la optimización multiobjetivo por lo general las funciones están en conflicto, así que es imposible (sin información adicional) seleccionar una única solución que optimice el problema (4.2). En estos casos se acepta un conjunto de soluciones, las cuales son consideradas igualmente válidas para el problema en cuestión. Llamaremos a este el conjunto de soluciones de Pareto, soluciones eficientes u óptimas de Pareto, y a su imagen vía F , el conjunto de puntos no dominados.

Definición 4.3.1. (Ehrgott, 2005, Definición 2.1). Decimos que un punto $\hat{x} \in \mathbb{R}^n$ es una solución de Pareto (también conocida como solución eficiente u óptimo de

⁴Un orden parcial se caracteriza por satisfacer las propiedades de reflexividad, transitividad y antisimetría - ver Definición 4.1.4-ítem 1

Pareto) del POM (4.2) ⁵, si no existe otra $x \in X$ tal que $F(x) \leq F(\hat{x})$, donde “ \leq ” denota el orden de Pareto definido en el espacio \mathbb{R}^p (ver Ejemplo 4.2.4).

Observación 4.3.1. Teniendo en cuenta la definición anterior, podemos decir que:

- Si \hat{x} es eficiente, entonces $F(\hat{x})$ es llamado **punto no dominado**.
- Si $x^1, x^2 \in X$ y $F(x^1) \leq F(x^2)$, entonces x^1 domina a x^2 (denotado por $x^1 \preceq x^2$) y $F(x^1)$ domina a $F(x^2)$.
- El conjunto de todas las soluciones eficientes (o soluciones de Pareto), denotado por \mathbf{X}_E , es llamado **conjunto eficiente** o variedad de Pareto.
- El conjunto de todos los puntos no dominados $\hat{y} = F(\hat{x}) \in Y$, donde $\hat{x} \in X_E$, denotado por $\mathbf{Y}_N = F(X_E)$, es llamado el **conjunto no dominado** o **frontera de Pareto**.

Se debe remarcar que las notaciones presentadas en las definiciones anteriores no son las únicas, en la literatura algunos autores llaman óptimo de Pareto, solución no dominada, solución no inferior, etc. a lo que nosotros estamos llamando solución de Pareto. Por ello, para evitar confusiones con la terminología, en la presente tesis usaremos las notaciones dadas por la Definición 4.3.1.

Entre las muchas definiciones utilizadas, decimos que \hat{x} es una solución de Pareto (o solución eficiente), si:

1. no existe $x \in X$ tal que $F_i(x) \leq F_i(\hat{x})$ para todo $i = 1, \dots, p$ y $F_j(x) < F_j(\hat{x})$ para al menos un índice $j \in \{1, \dots, p\}$;
2. no existe $x \in X$ tal que $F(x) - F(\hat{x}) \in -\mathbb{R}_{\geq}^p \setminus \{0\}$;
3. $F(x) - F(\hat{x}) \in \mathbb{R}^p \setminus \{-\mathbb{R}_{\geq}^p \setminus \{0\}\}$ para todo $x \in X$;
4. $Y \cap (F(\hat{x}) - \mathbb{R}_{\geq}^p) = \{F(\hat{x})\}$, donde $Y = F(X)$.

La definición 1, es una consecuencia inmediata de la Definición 4.3.1, en la cuál se utiliza el orden de Pareto “ \leq ” (detallado en el Ejemplo 4.2.4). La definición 2 también es una consecuencia de la Definición 4.3.1, la cuál indica que \hat{x} es una solución eficiente, si

$$\begin{aligned} \nexists x \in X : F(x) \leq F(\hat{x}) &\iff F(x) - F(\hat{x}) \leq 0 \\ &\iff F(x) - F(\hat{x}) \in -\mathbb{R}_{\geq}^p = -\mathbb{R}_{\geq}^p \setminus \{0\}, \end{aligned}$$

de donde se sigue dicho resultado, del cuál inmediatamente se deduce la definición 3. Finalmente la definición 4 se desprende de la definición 2. En efecto, supongamos que existe $x \in X$ tal que $F(x) \in Y \cap (F(\hat{x}) - \mathbb{R}_{\geq}^p)$, con $F(x) \neq F(\hat{x})$

⁵Si consideramos el POM (4.1), la solución $\hat{x} \in \mathbb{R}^n$ debe ser factible, es decir, debe satisfacer todas las restricciones de igualdad y/o desigualdad que componen al conjunto X .

entonces en particular existe un $d \in \mathbb{R}_{\geq}^p$, $d \neq 0$ tal que $F(x) = F(\hat{x}) - d$ ⁶, es decir $F(\hat{x}) - F(x) = d \in \mathbb{R}_{\geq}^p \setminus \{0\}$, lo cual implica que \hat{x} no sería eficiente, contradiciendo de esta manera la definición 2.

Las definiciones anteriores pueden ser ilustradas gráficamente. En la Figura 4.4 (b) se muestra una definición equivalente a 2. y 3.; mientras que en la Figura 4.4 (a), además de mostrar que las definiciones 1. y 4. son equivalentes a la Definición 4.3.1, se muestra también que $F(\hat{x})$ es un punto no dominado, dado que no existe $x \in X$, $x \neq \hat{x}$ tal que $F(x)$ pertenezca a la región $F(\hat{x}) - \mathbb{R}_{\geq}^p$, la cuál gráficamente es obtenida considerando el cuadrante negativo del plano cartesiano trazado a partir del punto $F(\hat{x})$.

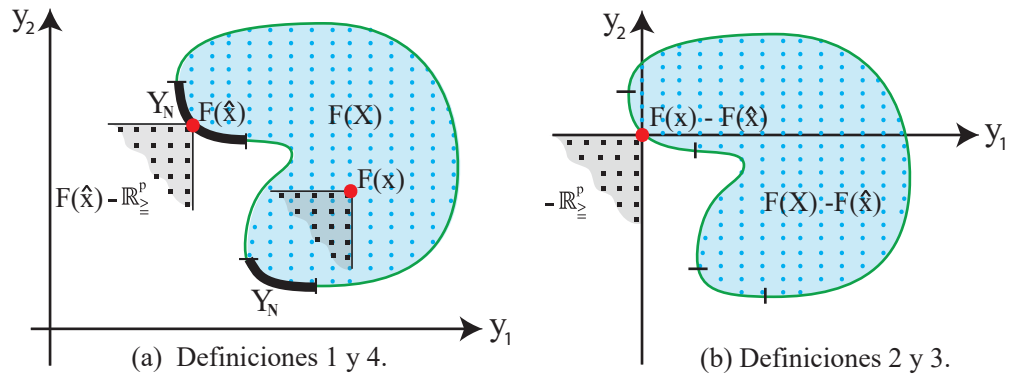


Figura 4.4: Ilustración de las definiciones de soluciones eficientes.

Con el objeto de aclarar los criterios de dominancia presentados en la Observación 4.3.1, a continuación presentamos su definición correspondiente, así como también algunos ejemplos que ilustrarán mejor dichas ideas.

Definición 4.3.2. Dominancia de Pareto (ver Sawaragi et al., 1985). Sean dos soluciones $x^1, x^2 \in \mathbb{R}^p$, se dice que x^1 domina (o es preferible) a x^2 (denotado por $x^1 \preceq x^2$), si $F(x^1) \leq F(x^2)$, donde “ \leq ” denota el orden de Pareto definido en el espacio \mathbb{R}^p (ver Ejemplo 4.2.4). En otras palabras x^1 domina a x^2 , si se cumplen las siguientes condiciones.

1. La solución x^1 es mejor o igual que x^2 en cada uno de los objetivos, es decir $F_i(x^1) \leq F_i(x^2)$ para todo $i = 1, \dots, p$.
2. La solución x^1 es estrictamente mejor que x^2 en al menos un objetivo, es decir $F_j(x^1) < F_j(x^2)$, para al menos un $j \in \{1, 2, \dots, p\}$.

Si cualesquiera de las condiciones es violada se dice que la solución x^1 no domina a la solución x^2 , en este caso denotamos $x^1 \not\preceq x^2$.

Ejemplo 4.3.1. Para facilitar la comprensión de los conceptos anteriores, vamos a abordar el POM tratado en el Ejemplo 4.0.1, pero considerando ahora solamente

⁶Esto se deduce del hecho que $F(x) \in (F(\hat{x}) - \mathbb{R}_{\geq}^p)$, donde $F(x) \neq F(\hat{x})$.

dos funciones objetivo F_1 y F_2 (precio y consumo). Las características de cada modelo de automóvil se muestran en la siguiente tabla:

Tabla 4.4: Características de los cuatro automóviles de la variante bi-objetivo del Ejemplo 4.0.1 (caso discreto).

Criterios de compra	Alternativas			
	VW	Opel	Ford	Toyota
Precio (1000 Dólares)	16.2	14.9	14.0	14.9
Consumo (l /100km)	7.2	7.0	7.5	8.2

La variación que estamos realizando al Ejemplo 4.0.1, nos permite obtener un problema de optimización bi-objetivo factible a ser ilustrado en un sistema de coordenadas bidimensional. En la Figura 4.5, las **regiones de dominancia**⁷, nos permiten identificar fácilmente, que las alternativas:

a) VW y Toyota pueden ser desechados, pues por ejemplo, Opel es más económico, y además presenta menor consumo que ambos (es preferible a ambos). Es decir Opel **domina** a Toyota y a VW, y éstos dos últimos son dominadas por Opel.

b) Opel y Ford son las soluciones a destacar (**puntos no dominados**), pues Opel consume menos combustible que los demás, y Ford es el más económico (para ambos no hay alternativas más baratas y que consuman menos petróleo).

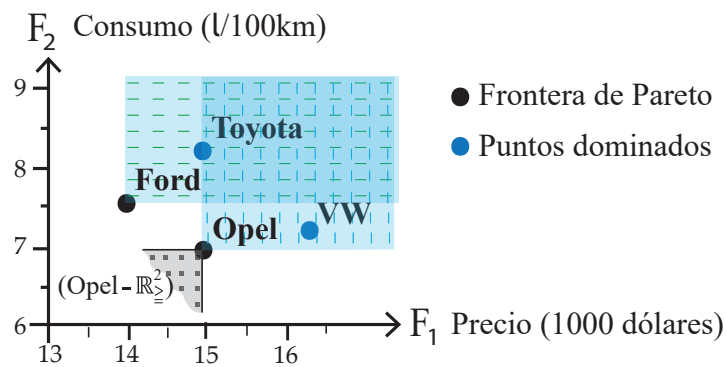


Figura 4.5: Espacio objetivo del Ejemplo 4.3.1.

Como no podemos escoger una única solución que minimice simultáneamente ambos objetivos (F_1 y F_2), entonces decimos que tanto Ford como Opel, son soluciones igualmente válidas para nuestro POM, y de acuerdo a la Definición 4.3.1, las mencionadas soluciones conforman la **frontera de Pareto**, es decir $Y_N = \{Ford, Opel\}$. \square

⁷Una región de dominancia, es el espacio que determinan las proyecciones positivas (en el caso de minimización) de dos semirrectas, trazadas a partir de un punto fijo. Por ejemplo, si en la Figura 4.5 consideramos el punto fijo Opel, su región de dominancia, es Toyota y VW.

Observación 4.3.2. Haciendo uso de la Definición 4.3.2, y considerando el conjunto $X = \{VW, Opel, Ford, Toyota\}$ y la función $F : X \rightarrow \mathbb{R}^2$, definida por $F(x) = (F_1(x), f_2(x)) = (\text{Precio}(x), \text{Consumo}(x))$, a continuación justificamos las consideraciones del Ejemplo (4.3.1).

a. Comparando el automóvil Opel y Ford con Toyota y VW tenemos que:

$Opel \preceq Toyota$, dado que $(14.9, 7.0) = F(Opel) \leq F(Toyota) = (14.9, 8.2)$.

La preferencia anterior se justifica por el hecho que $F_i(Opel) \leq F_i(Toyota)$ para todo $i = 1, 2$, y además existe al menos un $j = 2 \in \{1, 2\}$ tal que $7.0 = F_2(Opel) < F_2(Toyota) = 8.2$. Por lo tanto Opel es preferible a Toyota ($Opel \preceq Toyota$). Además, Opel pertenece a la frontera de Pareto, puesto que por la definición 4. (ver Observación 4.3.1) no existe otro automóvil en la región $Opel - \mathbb{R}_{\geq}^p$. Con semejantes consideraciones también tenemos que:

$Opel \preceq VW$, dado que $(14.9, 7.0) = F(Opel) \leq F(VW) = (16.2, 7.2)$.

$Ford \preceq Toyota$, dado que $(14.0, 7.5) = F(Ford) \leq F(Toyota) = (15.2, 8.2)$

$Ford \not\preceq VW$, dado que $(14.0, 7.5) = F(Ford) \not\leq F(VW) = (16.2, 7.2)$.

b. Finalmente comparando el automóvil Opel con Ford tenemos que:

$Opel \not\preceq Ford$, dado que $(14.9, 7.0) = F(Opel) \not\leq F(Ford) = (14.0, 7.5)$;

$Ford \not\preceq Opel$, dado que $(14.0, 7.5) = F(Ford) \not\leq F(Opel) = (14.9, 7.0)$.

Dado que por el ítem (a.) Ford y Opel representan los únicos puntos no dominados del problema de optimización bi-objetivo, entonces éstos puntos conforman la **frontera de Pareto**, representada por $Y_N = \{Ford, Opel\} = F(X_E)$, de donde inmediatamente se desprende que el **conjunto eficiente** $X_E = \{(14.9, 7.0); (14.0, 7.5)\}$. Finalmente por el ítem (b.) se tiene que las soluciones Opel y Ford son no dominadas entre sí.

Ejemplo 4.3.2. Continuemos con el análisis del Ejemplo 4.0.2, dado por

$$\min_{x \in \mathbb{R}} F(x) = \min_{x \in \mathbb{R}} (e^x, (e^x - 3)^2 + 1).$$

Anteriormente vimos que es fácil encontrar los minimizadores para cada una de las funciones objetivo, quedando pendiente la interrogante **¿cómo calcular el minimizador o los minimizadores del problema bi-objetivo?**. Para responder a ello, el proceso de minimización lo realizamos obteniendo la imagen del conjunto factible $X = \mathbb{R}$ (definido en el espacio de decisión), es decir se debe calcular el conjunto $F(X) = Y$ (el cuál se encuentra definido en el espacio objetivo); para ello realizamos lo siguiente:

$$(F_1(x), F_2(x)) = (y_1, y_2).$$

Dado que $y_1 = e^x$, entonces $y_2 = (e^x - 3)^2 + 1 = (y_1 - 3)^2 + 1$. La gráfica de ésta función, se muestra en la Figura 4.6 (a).

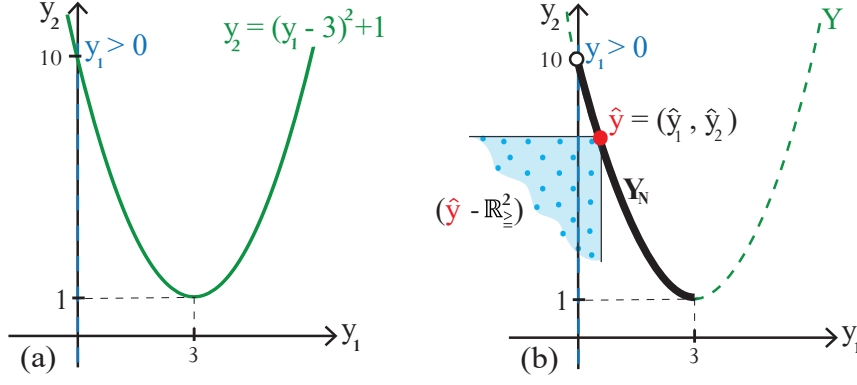


Figura 4.6: (a) Espacio objetivo y (b) puntos no dominados del Ejemplo 4.3.2.

Puesto que $y_1 = e^x > 0$, en la Figura 4.6 (b), se puede observar que el **conjunto no dominado o frontera de Pareto** Y_N se encuentran entre los correspondientes valores de $y_1 \in (0, 3]$ y $y_2 \in [1, 10]$; así que $Y_N = \{(y_1, y_2) \in Y : y_1 \in (0, 3]\}$.

Teniendo identificado Y_N , fácilmente se comprueba que el **conjunto de soluciones de Pareto** (o conjunto de soluciones eficientes) X_E está conformado por los $x \in (-\infty, \ln(3)]$ ⁸; así que $X_E = \{x \in X : (F_1(x), F_2(x)) \in Y_N\} = (-\infty, \ln(3)]$. En la Figura 4.6, también se puede observar que el ángulo recto adjunto al punto $\hat{y} = (\hat{y}_1, \hat{y}_2)$ muestra que no existe otro punto $y = (y_1, y_2) \in F(X)$, $y \neq \hat{y}$ tal que $y_1 \leq \hat{y}_1$ y $y_2 \leq \hat{y}_2$. \square

Observación 4.3.3. Haciendo uso de la Definición 4.3.2, y considerando los puntos $P_1(0.01, 9.9401)$, $P_2(1, 5)$, $P_3(3, 1)$, $P_4(5, 5)$, $P_5(5.99, 9.9401)$ y $P_6(7, 17)$, a continuación justificamos la no dominancia de algunos elementos pertenecientes al conjunto Y_N , calculado en el ejemplo anterior.

a. Comparando el punto P_1 con P_2 , P_3 , P_4 , P_5 y P_6 , tenemos que:

$P_1(0.01, 9.9401) \not\leq P_2(1, 5)$,	dado que al comparar componente a componente $0.01 \not\leq 1$ y $9.9401 \not\leq 5$;
$P_1(0.01, 9.9401) \not\leq P_3(3, 1)$,	pues a pesar de que $0.01 \leq 3$, se tiene que $9.9401 \not\leq 1$. Análogamente se tiene que $P_1 \not\leq P_4$;
$P_1(0.01, 9.9401) \preceq P_5(5.99, 9.9401)$,	dado que $0.01 < 5.99$ y $9.9401 \leq 9.9401$;
$P_1(0.01, 9.9401) \preceq P_6(7, 17)$,	pues $0.01 < 7$ y $9.9401 < 17$;

Los cálculos anteriores pueden ser simplificados, considerando una interpretación geométrica (ver Ejemplo 4.3.1), la cuál indica que si fijamos por ejemplo el punto P_1 , y trazamos a partir de él un sistema de coordenadas cartesianas, como se ve en la Figura 4.7, el cuadrante positivo que éste de-

⁸Como $y_1 \in (0, 3]$, entonces $y_1 = e^x \leq 3$. Así que $x \in (-\infty, \ln(3)]$.

termina, va a determinar la región de dominancia de P_1 , la cual está dada por el conjunto $(P_1 + \mathbb{R}_{\geq}^2)$; y el cuadrante negativo, va a determinar la región o el conjunto de todos los puntos que dominan a P_1 , la cuál esta dada por $(P_1 - \mathbb{R}_{\geq}^2)$.

Dado que $P_5, P_6 \in (P_1 + \mathbb{R}_{\geq}^2)$, entonces podemos decir que el punto P_1 domina a P_5 y P_6 (sin embargo no domina, por ejemplo a los puntos P_2, P_3 y P_4 , verificándose de esta manera los cálculos anteriores). Además, P_1 es un punto no dominado, esto es $P_1 \in Y_N$, dado que $Y \cap (P_1 - \mathbb{R}_{\geq}^2) = \{\emptyset\}$ (ver Observación 4.3.1 - ítem 4 de las definiciones alternativas de solución de Pareto).

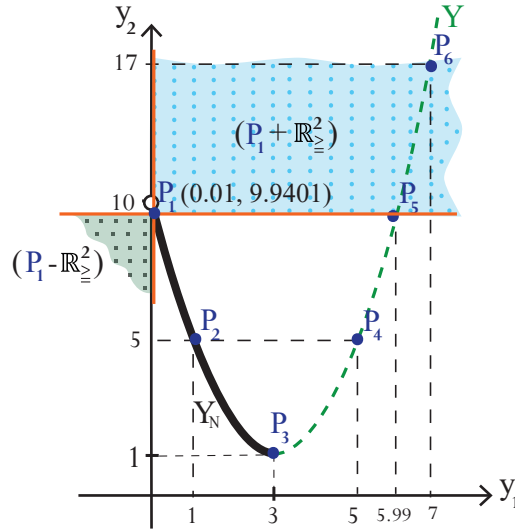


Figura 4.7: Regiones de dominancia del punto $P_1(0, 4)$.

b. Comparando el punto P_5 con P_2 y P_6 , tenemos que:

$P_5(5.99, 9.9401) \not\preceq P_2(1, 5)$, dado que $P_2 \notin (P_5 + \mathbb{R}_{\geq}^2)$;
 $P_5(5.99, 9.9401) \preceq P_6(9.9401, 17)$, dado que $P_6 \in (P_5 + \mathbb{R}_{\geq}^2)$. Sin embargo $P_5 \notin Y_N$, puesto que por ejemplo existen los puntos $P_i \in Y$, $i = 1, 2, 3, 4$, tales que $P_i \in (P_5 - \mathbb{R}_{\geq}^2)$.

Ejemplo 4.3.3. (Tablada, 2011, Ejemplo 6). Consideremos el siguiente problema de optimización multiobjetivo lineal:

$$\begin{cases} \min (2x_1 - x_2, -x_1 + 5x_2) \\ \text{s.a : } & x_1 + x_2 \leq 8 \\ & -x_1 + x_2 \leq 3 \\ & 0 \leq x_1 \leq 6 \\ & 0 \leq x_2 \leq 4. \end{cases}$$

El conjunto factible X (en el espacio de decisión) se muestra en la Figura 4.8 (a), y el conjunto factible Y (en el espacio objetivo) junto con el conjunto no dominado

Y_N se muestran en la Figura 4.8 (b) . El conjunto Y se obtienen tomando las imágenes de los vértices del conjunto X , como se observa en las siguiente tabla.

Tabla 4.7: Vértices del conjunto factible X en el espacio variable, e imágenes de los mismos en el espacio objetivo.

Vértices	x_1	x_2	Vértices	Y_1	Y_2
A	0	0	A'	0	0
B	0	3	B'	-3	15
C	1	4	C'	-2	19
D	4	4	D'	4	16
E	6	2	E'	10	4
F	6	0	F'	12	-6

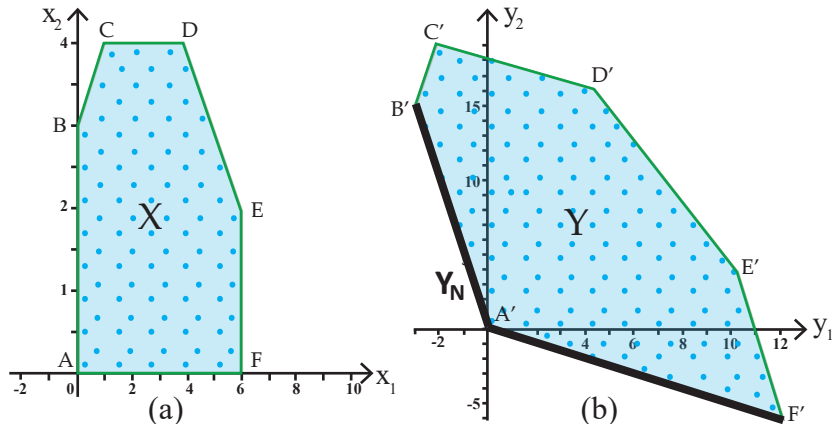


Figura 4.8: (a) Conjunto factible X en el espacio variable. (b) Conjunto factible Y y conjunto no dominado Y_N en el espacio objetivo.

Otros ejemplos de optimización multiobjetivo, así como papers, tesis, software y funciones de testeo con sus respectivos conjuntos no dominados (frentes de Pareto) pueden ser consultados en la siguiente página web: <http://delta.cs.cinvestav.mx/~ccoello/EMOO/>. \square

En lo que respecta al orden de Pareto, resulta de utilidad definir un orden en \mathbb{R}^p más restrictivo, que nos lleve a un concepto más débil de solución eficiente. Análogamente a lo realizado en el Ejemplo 4.2.4, considerando ahora el cono $C = \mathbb{R}_{>}^p = \{y \in \mathbb{R}^p : y_i > 0, \forall i = 1, \dots, p\}$ (ver relación (4.7)), definamos el orden “ $<$ ” en \mathbb{R}^p (ver la Tabla 4.2), por:

$$\begin{aligned}
 y^1 < y^2 &\iff y^2 - y^1 \in \mathbb{R}_{>}^p \\
 &\iff y_i^1 < y_i^2 \quad \forall k = 1, \dots, p.
 \end{aligned} \tag{4.12}$$

Teniendo en cuenta este orden presentamos la siguiente definición.

Definición 4.3.3. (Ehrgott, 2005, Definición 2.24). Un punto $\hat{x} \in \mathbb{R}^n$ es una solución débil de Pareto (llamada también solución débilmente eficiente o débilmente óptima de Pareto) del POM (4.2) ⁹, si no existe $x \in X$ tal que $F(x) <^{10} F(\hat{x})$, es decir $F_i(x) < F_i(\hat{x})$ para todo $i = 1, \dots, p$.

Teniendo en cuenta la definición anterior podemos decir que:

- Si \hat{x} es una solución débilmente eficiente, entonces el punto $F(\hat{x}) = \hat{y}$ es llamado **débilmente no dominado**.
- La **débil eficiencia y no dominancia** de conjuntos, son denotados por X_{wE} y Y_{wN} , respectivamente.

Algunos autores consideran que un punto débilmente no dominado es un punto no dominado respecto al $\text{int}(\mathbb{R}_{\leq}^p) = \mathbb{R}_{>}^p$, notación que es bastante conveniente en el contexto de los conos de eficiencia y de no dominancia.

De las Definiciones 4.3.1 y 4.3.3 se sigue que:

$$Y_N \subset Y_{wN} \quad (4.13)$$

$$X_E \subset X_{wE} \quad (4.14)$$

Además, como en el caso de eficiencia, la débil eficiencia tiene otras definiciones equivalentes. Se dice que una solución factible $\hat{x} \in X \subseteq \mathbb{R}^n$ es débilmente eficiente, si:

1. no existe un $x \in X$ tal que $F(\hat{x}) - F(x) \in \text{int}(\mathbb{R}_{\leq}^p) = \mathbb{R}_{>}^p$;
2. $(F(\hat{x}) - \mathbb{R}_{>}^p) \cap Y = \emptyset$.

De las definiciones anteriores se desprende que:

$$Y_{wN} = \{\hat{y} \in Y : \nexists y \in Y, y < \hat{y}\} = \{\hat{y} \in Y : (\hat{y} - \mathbb{R}_{>}^p) \cap Y = \emptyset\}. \quad (4.15)$$

Ejemplo 4.3.4. Consideremos el siguiente problema de optimización multiobjetivo:

$$\begin{cases} \min F(x) = (x_1, x_2) \\ \text{s.a. : } 0 < x_1 < 1 \\ 0 \leq x_2 \leq 1. \end{cases}$$

De forma análoga al cálculo realizado en el Ejemplo 4.3.3, el conjunto factible Y (definido en el espacio de decisión) está dado por

$$Y = \{(y_1, y_2) \in \mathbb{R}^2 : 0 < y_1 < 1; 0 \leq y_2 \leq 1\}, \quad (4.16)$$

⁹ Si consideramos el POM (4.1), la solución $\hat{x} \in X$ debe ser factible, es decir, debe satisfacer todas las restricciones de igualdad y/o desigualdad que componen al conjunto X .

¹⁰ “ $<$ ” denota el orden estricto de Pareto definido en (4.12).

En base al conjunto anterior, como se puede ver en la Figura 4.9 (a), se tiene que $Y_N = \emptyset$, y $Y_{wN} = (0, 1) \times \{0\} = \{y \in Y : 0 < y_1 < y_2, y_2 = 0\}$.

Modificando las restricciones del problema, de tal manera que el conjunto (4.16) sea un cuadro cerrado (ver Figura 4.9 (b)), esto es:

$$Y = \{(y_1, y_2) \in \mathbb{R}^2 : 0 \leq y_i \leq 1\}, \quad (4.17)$$

se tiene que $Y_N = \{0\}$ y $Y_{wN} = \{(y_1, y_2) \in Y : y_1 = 0 \text{ o } y_2 = 0\}$. \square

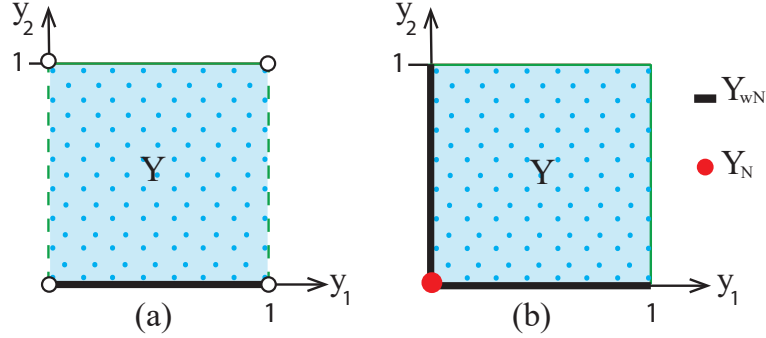


Figura 4.9: (a) Y_N es vacío, pero Y_{wN} no lo es. (b) Punto no dominado, y puntos débilmente no dominados.

Observación 4.3.4. De forma análoga a las observaciones 4.3.3 y 4.3.3, considerando los puntos $P_1(0, 0)$, $P_2(0, \frac{1}{2})$, $P_3(\frac{1}{4}, 0)$, $P_4(\frac{1}{2}, 0)$, $P_5(\frac{1}{2}, \frac{1}{2})$ y $P_6(\frac{3}{4}, \frac{1}{2})$ pertenecientes (o no), a los conjuntos representados en la Figura 4.9, a continuación mostramos la eficiencia, y débil eficiencia de los mismos.

a) Consideremos el conjunto Y representado en la Figura 4.10 (a).

a.1. Comparando el punto P_3 con P_5 , tenemos que:

$$P_3(\frac{1}{4}, 0) \prec P_5(\frac{1}{2}, \frac{1}{2}), \text{ dado que al comparar componente a componente se tiene que } \frac{1}{4} < \frac{1}{2} \text{ y } 0 < \frac{1}{2}. \text{ Además, } P_3 \in Y_{wN} \text{ dado que } (P_3 - \mathbb{R}_{>}^2) \cap Y = \emptyset.$$

a.2. Comparando el punto P_5 con P_6 , tenemos que:

$$P_5(\frac{1}{2}, \frac{1}{2}) \preceq P_6(\frac{3}{4}, \frac{1}{2}), \text{ dado que } \frac{1}{2} \leq \frac{3}{4} \text{ y } \frac{1}{2} \leq \frac{1}{2}. \text{ Es decir } P_5 \in (P_6 - \mathbb{R}_{\geq}^2); \text{ sin embargo } P_5 \notin Y_{wN}, \text{ dado que por lo menos existen los puntos } P_1, P_3 \in (P_5 - \mathbb{R}_{>}^2).$$

b) Consideremos el conjunto Y representado en la Figura 4.10 (b).

b.1. Comparando el punto P_1 con P_2 , se tiene que

$$P_1(0,0) \preceq P_2(0,\frac{1}{2}), \text{ dado que } 0 \leq 0 \text{ y } 0 \leq \frac{1}{2}.$$

Como se puede ver en la Figura 4.10 $P_1 \in Y$ es el único punto no dominado, dado que $(P_1 - \mathbb{R}_{\geq}^2) \cap Y = \emptyset$.

b.2. Comparando P_4 con los puntos P_5 y P_6 se puede ver que a parte de dominar a ambos (pues $P_5, P_6 \in (P_4 + \mathbb{R}_{\geq}^2)$), P_4 es un punto débilmente no dominado, es decir $P_4 \in Y_{wN}$, dado que $(P_4 - \mathbb{R}_{>}^2) \cap Y = \emptyset$.

b.3. Considerando como fijo el punto P_5 , se puede ver que los puntos $P_1, P_3 \in Y$ lo dominan, dado que $P_1, P_3 \in (P_5 - \mathbb{R}_{\geq}^2)$.

b.4. Considerando como fijo el punto P_4 , se puede ver que $P_4 \notin Y_N$, dado que por ejemplo existe el punto $P_3 \in (P_4 - \mathbb{R}_{\geq}^2)$. Sin embargo P_4 es débilmente eficiente, es decir $P_4 \in Y_{wN}$, dado que $(P_4 - \mathbb{R}_{>}^2) \cap Y = \emptyset$.

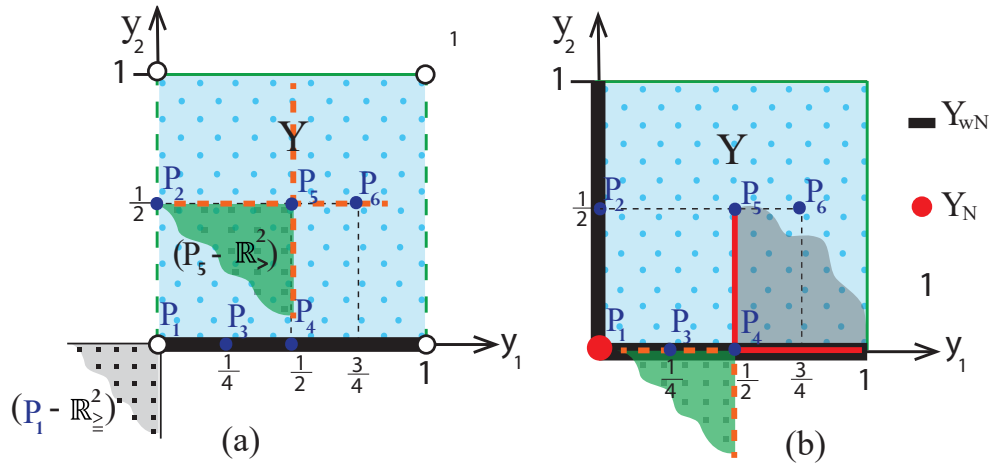


Figura 4.10: Regiones de dominancia de los conjuntos Y_N y Y_{wN} .

El siguiente resultado que nos indican que los puntos eficientes están incluidos en la frontera de Y .

Proposición 4.3.1. $Y_N \subset Fr(Y)$.

Demostración. Sea $y \in Y_N$, y supongamos que $y \notin Fr(Y)$, entonces $y \in int(Y)$, luego existe una bola $B(y, \epsilon)$ tal que $B(y, \epsilon) = (y + B(0, \epsilon)) \subset Y$, donde $B(0, \epsilon)$ es una bola abierta con radio ϵ y centro el origen. Sea $d \in \mathbb{R}_{\geq}^p$, $d \neq 0$, entonces podemos escoger algún $\alpha \in \mathbb{R}$, $0 < \alpha < \epsilon$ tal que $\alpha d \in B(0, \epsilon)$. Luego $(y \pm \alpha d) \in Y$, con $\alpha d \in \mathbb{R}_{\geq}^p \setminus \{0\}$, lo cual implica que $y \notin Y_N$ (pues $(y - \alpha d) \in Y$). \square

Capítulo 5

MÉTODO DE PUNTO PROXIMAL ESCALARIZADO INEXACTO (MPPEI)

5.1. Antecedentes de la investigación

El Método del Punto Proximal (MPP), introducido por Martinet (1970), y posteriormente utilizado por Rockafellar (1976) para encontrar ceros de operadores monótonos maximales, surgió inicialmente para resolver el problema

$$\min \{f(x) : x \in \mathbb{R}^n\}, \quad (5.1)$$

donde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es una función convexa. El algoritmo de punto proximal genera una sucesión $\{x^k\}$, a través del siguiente proceso iterativo: dado un punto inicial arbitrario $x^0 \in \mathbb{R}^n$, calcular

$$x^{k+1} = \operatorname{argmin}\{f(x) + \frac{\lambda_k}{2}\|x - x^k\|^2, x \in \mathbb{R}^n\}, \quad (5.2)$$

donde $\lambda_k > 0$ es el parámetro de regularización, $\|\cdot\|$ denota la norma Euclidiana, y la notación argmin^1 significa que x^k es el único punto mínimo global de $f(x) + \frac{\lambda_k}{2}\|x - x^k\|^2$ en \mathbb{R}^n . Si $\{\lambda_k\}$ satisface $\sum_{k=1}^{+\infty} (1/\lambda_k) < +\infty$, se demuestra que la sucesión $\{f(x^k)\}$ converge al ínfimo de f ; y si además el conjunto de soluciones óptimas es no vacío, entonces $\{x^k\}$ converge a una solución óptima del problema, ver Güler (1992).

*Observar que cuando la función f definida en (5.1) no es convexa, resolver el subproblema (5.2) podría ser incluso más difícil que resolver el problema original. Para superar tal inconveniente, en el año 2012 Papa Quiroz y Oliveira presentan el artículo “**An extension of proximal methods for quasiconvex minimization on the nonnegative orthant**” (Papa Quiroz y Oliveira, 2012), el cuál tiene como objetivo resolver el problema (5.1), donde ahora la función $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ es propia, semicontinua inferior y cuasi-convexa, y $X = \{x \in \mathbb{R}^n : x \geq 0\}$ es un conjunto convexo abierto no vacío. Los autores tomando un punto inicial $x^0 \in \operatorname{dom}(f) \cap \mathbb{R}^n$, generan una sucesión $\{x^k\}$ que converge a una solución del problema. Para tal fin en vez de utilizar la iteración (5.2), propusieron*

$$0 \in \hat{\partial} (f(\cdot) + \lambda_k d(\cdot, x^{k-1})) (x^k), \quad (5.3)$$

¹Dada la función $g : \mathbb{R}^n \supset X \rightarrow \mathbb{R}$, decimos que $z \in \operatorname{argmin}\{g(y) : y \in X\}$, si $g(z) \leq g(y)$ para todo $y \in X$.

donde $\hat{\partial}$ es el subdiferencial de Fréchet, y d es una distancia proximal introducida por Auslender y Teboulle (2006). Ellos probaron que la sucesión $\{x^k\}$ está bien definida, y se encuentra en el ortante positivo. Asumiendo que f es cuasi-convexa, prueban que la sucesión generada por el algoritmo converge hacia alguna solución óptima del problema, asumiendo también que los parámetros λ_k son acotado y f es continuamente diferenciable, prueban que tal sucesión converge a un punto KKT.

Con el objetivo de construir algoritmos de punto proximal computacionalmente más implementables, que resuelvan problemas cuasi-convexos, dichos algoritmos han sido extendidos hacia sus versiones inexactas, lo cual se justifica puesto que por lo general es muy difícil encontrar un valor x^k que resuelva de forma exacta, por ejemplo el subproblema de optimización original (5.2); es así que se consideran errores de aproximación e^k , los cuales si cumplen ciertos requisitos, generalizan las propiedades de convergencia de los algoritmos de búsqueda exacta. Por ejemplo Papa Quiroz et al. (2015), en respuesta a ciertas interrogantes planteadas en un trabajo previo Papa Quiroz y Oliveira (2012), presentan el artículo “**An inexact proximal method for quasiconvex minimization**”, el cual tiene como objeto resolver el problema (5.1), donde f satisface las mismas condiciones que el problema anterior, pero ahora se encuentra restringido a la cerradura de un conjunto abierto arbitrario X . Para probar que la sucesión $\{f(x^k)\}$ es decreciente, los autores en lugar de (5.3), propusieron la siguiente iteración inexacta:

$$\|x^{k-1} - x^k - e^{k+1}\| \leq \max\{\|e^k\|, \|x^k - x^{k-1}\|\}, \quad (5.4)$$

donde d es una distancia proximal que satisface ciertas condiciones relacionadas a las distancias proximales inducidas, y e^{k+1} es el error de aproximación, el cual para asegurar que la sucesión $\{x^k\}$ cumpla las propiedades de Cuasi-Fejér convergencia, satisface las siguientes condiciones:

$$\sum_{k=1}^{+\infty} \frac{\|e^{k+1}\|}{\lambda_k} < +\infty \quad y \quad \sum_{k=1}^{+\infty} \frac{|\langle e^{k+1}, x^k \rangle|}{\lambda_k} < +\infty.$$

En lo referente al Método de Punto Proximal para minimización de problemas multiobjetivo, se puede citar a Bento et al. (2014), quienes presentan el Paper “**Proximal point method for a special class of nonconvex multiobjective optimization problem**”, el cuál tiene como objetivo resolver el problema (5.1) en su versión multibjetivo ($F : \mathbb{R}^n \rightarrow \mathbb{R}^p$) irrestricta, donde cada una de las funciones $F_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, \dots, p$, son diferenciables no necesariamente convexas. Los autores tomando como punto inicial x^0 , generan una sucesión de soluciones convergentes x^{k+1} . Para tal fin hacen uso de la siguiente iteración:

$$x^{k+1} \in \operatorname{argmin}_w \left\{ F(\cdot) + \frac{\lambda_k}{2} \|\cdot - x^k\|^2 : x \in \Omega_k \right\} \quad k = 0, 1, \dots, \quad (5.5)$$

donde $\lambda_k > 0$ es el parámetro de regularización, $\Omega_k = \{x \in \mathbb{R}^n : F(x) \preceq F(x^k)\}$, $\lambda_k > 0$, $\{z_k\} \subset \mathbb{R}_{\geq}^p \setminus \{0\}$, $\|z_k\| = 1$, $\|\cdot\|$ denota la norma Euclidiana, y la notación argmin_w significa que x^k es el único punto mínimo global débil de $F(x) + \frac{\lambda_k}{2} \|x - x^k\|^2$ en \mathbb{R}^n . Los autores prueban que la sucesión está bien definida,

y que en el caso convexo o pseudo-convexo cualquier punto de acumulación generado por la sucesión es un punto crítico Pareto-Clarke. Además, bajo algunos supuestos adicionales prueban la convergencia global del método.

Es así que la iteración original (5.3) ha ido variando en función de la naturaleza de los problemas de optimización, y del espacio en el cuál están definidos. En el año 2016 Apolinário, Papa Quiroz y Oliveira, trabajando ahora con funciones vectoriales $F_i : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, $i \in \{1, \dots, p\}$ cuasi-convexas, localmente Lipschitz, y acotadas interiormente por el cero, extienden el Método de Punto Proximal multiobjetivo presentado por Bento et al. (2014). Los autores presentan el artículo "**A scalarization proximal point method for quasiconvex multiobjective minimization**" (ver Apolinário et al., 2016), el cual tiene como objeto resolver el problema (5.1) en su versión multiobjetivo. La iteración que consideran en lugar de (5.5) es la siguiente. Dado un punto inicial arbitrario $x^k \in \mathbb{R}^n$, encontrar un $x^{k+1} \in \Omega_k$, tal que:

$$0 \in \partial^0 \left(\langle F(\cdot), z_k \rangle + \frac{\lambda_k}{2} \|\cdot - x^k\|^2 \right) (x^{k+1}) + \mathcal{N}_{\Omega_k}(x^{k+1}), \quad (5.6)$$

donde ∂^0 es el subdiferencial de Clarke, $\Omega_k = \{x \in \mathbb{R}^n : F(x) \preceq F(x^k)\}$, $\lambda_k > 0$, $\{z_k\} \subset \mathbb{R}_{\geq}^p \setminus \{0\}$, $\|z_k\| = 1$, y $N_{\Omega_k}(x^{k+1})$ es el cono normal a Ω_k en x^{k+1} . Con dicha iteración (la cual se detiene cuando $x^{k+1} = x^k$ o cuando x^{k+1} es un punto crítico Pareto-Clarke), prueban que la sucesión $\{x^{k+1}\}$ generada por el método está bien definida, y además obtienen la convergencia global a un punto crítico Pareto-Clarke, y cuando F es convexa, obtienen la convergencia a una solución débil de Pareto.

5.2. El algoritmo

Con el objetivo de resolver el problema de optimización multiobjetivo POM (4.2), dado por:

$$\min_{x \in \mathbb{R}^n} F(x) = \min\{F(x), x \in \mathbb{R}^n\}, \quad (5.7)$$

donde $F : \mathbb{R}^n \rightarrow \mathbb{R}^p$ satisface las siguientes hipótesis:

- (H1) $0 \preceq F$, es decir que $F_i(x) \geq 0$ para todo $i \in \{1, \dots, p\}$ y $x \in \text{dom}(F_i)$,
- (H2) F es localmente Lipschitz en \mathbb{R}^n (ver Definición 2.2.3 y 2.2.13),
- (H3) F es \mathbb{R}_{\geq}^p -cuasi-convexa (ver Definición 2.2.11).

Proponemos el Método de Punto Proximal Escalarizado Inexacto (MPPEI) con regularización cuadrática. Definido por:

Algoritmo MPPEI

Inicio: Escoger un punto inicial arbitrario

$$x^0 \in \mathbb{R}^n. \quad (5.8)$$

Paso iterativo: Para $k = 0, 1, \dots$, y dado un x^k , encontrar $x^{k+1} \in \Omega_k$ tal que:

$$e^{k+1} \in \partial^0 \left(\langle F(\cdot), z_k \rangle + \frac{\alpha_k}{2} \|\cdot - x^k\|^2 \right) (x^{k+1}) + \mathcal{N}_{\Omega_k}(x^{k+1}), \quad (5.9)$$

donde $\Omega_k = \{x \in \mathbb{R}^n : F(x) \preceq F(x^k)\}$, $\alpha_k > 0$, $\{z_k\} \subset \mathbb{R}_{\geq}^p \setminus \{0\}$ y $\|z_k\| = 1$.

Criterio de parada: Si $x^{k+1} = x^k$ o si x^{k+1} es un punto crítico Pareto-Clarke (ver Definición 3.2.1), detener el algoritmo. En otro caso hacer $k \leftarrow k + 1$ y retornar al paso iterativo.

Observación 5.2.1. 1. Si $e^{k+1} = 0$ para todo $k \in \mathbb{N}$, entonces en la expresión (5.9) estaremos ante la versión exacta del Método de Punto Proximal Escalarizado (MPPE), propuesto por Apolinário et al. (2016). Así se puede decir que el algoritmo que se está proponiendo representa una extensión hacia su versión inexacta del MPPE.

Además, si F es \mathbb{R}_{\geq}^p -convexa, entonces la expresión (5.9) resulta en

$$x^{k+1} = \operatorname{argmin} \left\{ \langle F(\cdot), z_k \rangle + \frac{\alpha_k}{2} \|\cdot - x^k\|^2 : x \in \Omega_k \right\},$$

iteración que ha sido estudiada por Bonnel et al. (2005), la cual fue tomada como referencia para la realización del MPPE (considerando en este caso funciones cuasi-convexas).

2. Si $F : \mathbb{R}^n \rightarrow \mathbb{R}$, $z_k = 1$, $\partial^0 = \widehat{\partial}$ (donde $\widehat{\partial}$ representa el subdiferencial de Fréchet), y $x^{k+1} \in \operatorname{int}(\Omega_k)$, entonces en (5.9) tendremos la siguiente iteración

$$e^{k+1} \in \widehat{\partial} \left(f(\cdot) + \frac{\alpha_k}{2} \|\cdot - x^k\|^2 \right) (x^{k+1}),$$

estudiada por Papa Quiroz et al. (2015). Además, si en la expresión anterior hacemos que $e^{k+1} = 0$, y consideramos algunas hipótesis adicionales, como que $x \geq 0$, obtendremos el algoritmo estudiado por Papa Quiroz y Oliveira (2012).

Observación 5.2.2. En la práctica para obtener el punto x^{k+1} que satisfaga (5.9), solamente se debería encontrar un punto crítico (mínimo local, máximo local o punto silla), y no necesariamente un mínimo del siguiente problema de optimización:

$$\min \left\{ \langle F(\cdot), z_k \rangle + \frac{\alpha_k}{2} \|\cdot - x^k\|^2 : x \in \Omega_k \right\}, \quad (5.10)$$

o equivalentemente, se debería encontrar un punto crítico del siguiente problema de optimización global con restricciones de desigualdad

$$\min \left\{ \sum_{i=1}^p z_k^i F_i(\cdot) + \frac{\alpha_k}{2} \|\cdot - x^k\|^2 : F_i(x) \leq F_i(x^k), \quad \forall i = 1, \dots, p \right\}. \quad (5.11)$$

5.2.1. Existencia de los iterados

Teorema 5.2.1. Si $F : \mathbb{R}^n \rightarrow \mathbb{R}^p$ es una función que satisface las hipótesis **(H1)**, **(H2)** y **(H3)**, entonces la sucesión $\{x^k\}$ generada por el algoritmo MPPEI está bien definida.

Demostración. Procedemos por inducción. Para $k = 0$, por (5.8) se tiene la existencia de x^0 . Asumamos que x^k existe, y definamos la función

$$\varphi_k(x) = \langle F(x), z_k \rangle + \frac{\alpha_k}{2} \|x - x^k\|^2 + \delta_{\Omega_k}(x),$$

donde $\delta_{\Omega_k}(\cdot)$ es la función indicadora o indicatriz de Ω_k (ver Definición 2.2.6), definida por

$$\delta_{\Omega_k}(x) = \begin{cases} 0 & \text{si } x \in \Omega_k \\ +\infty & \text{si } x \notin \Omega_k. \end{cases} \quad (5.12)$$

Así tenemos la siguiente equivalencia:

$$\min\{\varphi_k(x) : x \in \mathbb{R}^n\} \equiv \min\{\langle F(x), z_k \rangle + \frac{\alpha_k}{2} \|x - x^k\|^2 : x \in \Omega_k\}.$$

Analizando cada una de las funciones que componen $\varphi_k(x)$, tenemos que:

- $\langle F(x), z_k \rangle$ está acotada inferiormente, puesto que $0 \preceq F$ (ver **(H1)**) y $z_k \in \mathbb{R}_{\geq}^p \setminus \{0\}$; además es continua (por ende también será sci), dado que $\langle F(x), z_k \rangle$ es localmente Lipschitz, como probamos a continuación.

Dado que por **(H2)** F es localmente Lipschitz (ver las definiciones 2.2.3 y 2.2.13), entonces para todo $w \in \text{dom}(F)$ existen $\epsilon_w > 0$ y $L_w > 0$, tales que

$$|F_i(x) - F_i(y)| \leq L_w \|x - y\| \quad \forall x, y \in B(w, \epsilon_w) \text{ y } \forall i \in \{1, \dots, p\}, \quad (5.13)$$

donde L_w es la constante de Lipschitz. Dado que $z_k \in \mathbb{R}_{\geq}^p \setminus \{0\}$, por la desigualdad de Cauchy-Schwarz (ver Teorema 2.1.2) y la aplicación de (5.13) tenemos que

$$\begin{aligned} |\langle F(x), z \rangle - \langle F(y), z \rangle| &= |\langle F(x) - F(y), z \rangle| \leq \|F(x) - F(y)\| \|z\| \\ &\leq L_w \|z\| \|x - y\|, \end{aligned} \quad (5.14)$$

lo cual implica que $\langle F(\cdot), z_k \rangle$ es localmente Lipschitz, puesto que $L_w \|z\| > 0$.

- $\|x - x^k\|^2$ también es continua (por ello también será sci), está acotada inferiormente (pues $0 \leq \|x - x^k\|^2$) y además es coerciva (ver Ejemplo 2.2.1).
- $\delta_{\Omega_k}(x)$ por la Proposición 2.2.1 es sci, además se encuentra acotada inferiormente (esto se deduce de la expresión (5.12), la cuál indica que $\delta_{\Omega_k}(x) \geq 0$ para todo $x \in \mathbb{R}^n$).

Sumando las tres funciones anteriores se tiene que $\varphi_k(x)$ es una función propia (dado que las tres son propias), sci (pues la suma de dos o más funciones semicontinuas inferiores es sci) y coerciva (la suma de una función coerciva con otra(s) que sea(n) acotada(s) inferiormente, resulta una función coerciva). Así que por la Proposición 2.2.2 se tiene que el valor óptimo de la función $\varphi_k(x)$ existe, y es finito. Como dicho valor existe, lo denotamos por x^{k+1} , y como es finito se tiene que $x^{k+1} \in \Omega_k$ ².

Dado que la función $\varphi_k(x)$ es coerciva y sci, por el Teorema 2.3.2 (x^{k+1} representa un mínimo global de $\varphi_k(x)$) y la Proposición 3.1.2 se tiene que

$$0 \in \widehat{\partial} \left(\langle F(\cdot), z_k \rangle + \frac{\alpha_k}{2} \|\cdot - x^k\|^2 + \delta_{\Omega_k}(\cdot) \right) (x^{k+1}).$$

Usando la Proposición 3.1.1 la expresión anterior implica que

$$0 \in \partial^{Lim} \left(\langle F(\cdot), z_k \rangle + \frac{\alpha_k}{2} \|\cdot - x^{k-1}\|^2 + \delta_{\Omega_k}(\cdot) \right) (x^{k+1}).$$

Puesto que $\langle F(\cdot), z_k \rangle + \frac{\alpha_k}{2} \|\cdot - x^k\|^2$ es localmente Lipschitz y $\delta_{\Omega_k}(\cdot)$ es sci (ver Proposición 2.2.1), por la Proposición 3.1.4 en la expresión anterior se tiene que

$$0 \in \partial^{Lim} \left(\langle F(\cdot), z_k \rangle + \frac{\alpha_k}{2} \|\cdot - x^k\|^2 \right) (x^{k+1}) + \partial^{Lim}(\delta_{\Omega_k}(\cdot))(x^{k+1}).$$

Como Ω_k es un conjunto convexo cerrado y $x^{k+1} \in \Omega_k$, entonces por la Proposición 3.1.3 en la expresión anterior tenemos que

$$0 \in \partial^{Lim} \left(\langle F(\cdot), z_k \rangle + \frac{\alpha_k}{2} \|\cdot - x^k\|^2 \right) (x^{k+1}) + N_{\Omega_k}(x^{k+1}). \quad (5.15)$$

Finalmente haciendo uso de la Observación 3.1.1, y considerando $e^{k+1} = 0$, la iteración (5.9) es obtenida de (5.15). \square

Como estamos interesados en la convergencia asintótica del algoritmo MPPEI, a lo largo de este trabajo vamos a asumir que en cada iteración x^{k+1} no es un punto crítico Pareto-Clarke, lo cual en función de la Observación 3.2.1-relación (3.19) va a implicar que $\text{int}(\Omega_{k+1}) \neq \emptyset$. También vamos a asumir que (en cada iteración) $x^{k+1} \neq x^k$; si esta condición no es satisfecha, es decir si existe un k_0 tal que $x^{k_0+1} = x^{k_0}$, entonces de forma análoga a lo realizado en los Teoremas

²Dado que el valor óptimo de $\varphi_k(x)$ es finito, cada uno de los términos que la componen son finitos. Centrándonos en la función $\delta_{\Omega_k}(\cdot)$ (ver la relación (5.12)), por la definición de la misma la única posibilidad para que x^{k+1} sea finito, es que $x^{k+1} \in \Omega_k$, puesto que si $x^{k+1} \notin \Omega_k$ se tendría que este término sería infinito.

5.3.1 y 5.3.2 (considerando las propiedades que éstos requieren), fácilmente se puede probar que x^{k_0+1} es un punto crítico Pareto-Clarke de F .

A continuación presentamos otra hipótesis sobre la función F y el punto inicial x^0 , la cual está citada en varios trabajos concernientes a los algoritmos proximales. Ver por ejemplo Bonnel et al. (2005), Ceng y Yao (2007) y Villacorta y Oliveira (2011). El conjunto

$$(\mathbf{H4}) \ (F(x^0) - \mathbb{R}_{\leq}^p) \cap F(\mathbb{R}^n) \text{ es } \mathbb{R}_{\leq}^p\text{-completo,}$$

lo cual significa que para toda sucesión $\{a^k\} \subset \mathbb{R}^n$, con $a^0 = x^0$, tal que $F(a^{k+1}) \preceq F(a^k)$ para todo $k \in \mathbb{N}$, existe $a \in \mathbb{R}^n$ tal que $F(a) \preceq F(a^k)$, para todo $k \in \mathbb{N}$.

Denotemos el conjunto:

$$E = \{x \in \mathbb{R}^n : F(x) \preceq F(x^k), \forall k \in \mathbb{N}\}. \quad (5.16)$$

Dicho conjunto acompañado de las hipótesis $(\mathbf{H1}) - (\mathbf{H4})$ satisface lo siguiente:

- $E \subseteq \Omega_k$, para cualquier $k \in \mathbb{N}$,
- $E \neq \emptyset$, dado que por el Teorema 5.2.1 la sucesión $\{x^{k+1}\}$ generada por el MPPEI existe, y satisface $F(x^{k+1}) \preceq F(x^k)$. Entonces por la hipótesis $(\mathbf{H4})$ existe un $a \in \mathbb{R}^n$ tal que $F(a) \preceq F(x^k)$, así que $a \in E$.
- Además, por las hipótesis $(\mathbf{H2})$ y $(\mathbf{H3})$ es claro que dicho conjunto también es cerrado y convexo, respectivamente.

5.2.2. Resultados de convergencia

El siguiente resultado jugará un papel esencial en el desarrollo de la presente tesis. Más precisamente será útil en el análisis de convergencia de la sucesión generada por el algoritmo MPPEI.

Lema 5.2.1. Sea $F : \mathbb{R}^n \rightarrow \mathbb{R}^p$ una función que satisface las hipótesis $(\mathbf{H1})$, $(\mathbf{H2})$ y $(\mathbf{H4})$, y $\{x^k\}$ la sucesión generada por el algoritmo MPPEI. Si \hat{x} es un punto de acumulación de $\{x^k\}$, entonces $\hat{x} \in E$.

Demostración. Sea $\hat{x} \in \mathbb{R}^n$ un punto de acumulación de $\{x^k\}$, entonces por el Teorema 2.1.11, existe una subsucesión $\{x^{k_j}\} \subseteq \{x^k\}$ tal que $\lim_{j \rightarrow +\infty} x^{k_j} = \hat{x}$. Probaremos que $\hat{x} \in E$.

Dado que por la hipótesis $(\mathbf{H2})$ F es localmente Lipschitz en \mathbb{R}^n , la función $\langle F(\cdot), z \rangle$ también es localmente Lipschitz (ver (5.13) y (5.14)), y por lo tanto continua en \mathbb{R}^n para todo $z \in \mathbb{R}^p$, en particular para $z \in \mathbb{R}_{\leq}^p \setminus \{0\}$; y como $x^{k_j} \rightarrow \hat{x} \in \mathbb{R}^n$, por el Teorema 2.1.13 tenemos que

$$\lim_{j \rightarrow +\infty} \langle F(x^{k_j}), z \rangle = \langle F(\hat{x}), z \rangle.$$

Por otro lado, se conoce que $x^{k+1} \in \Omega_k$ (esto se demostró en el Teorema 5.2.1), entonces $F(x^{k+1}) \preceq F(x^k)$ ³, así que $F_i(x^{k+1}) \leq F_i(x^k)$ para todo $i \in \{1, \dots, p\}$; y como $z \in \mathbb{R}_{\geq}^p \setminus \{0\}$, entonces

$$\langle F(x^{k+1}), z \rangle \leq \langle F(x^k), z \rangle. \quad (5.17)$$

Dado que la función $\langle F(\cdot), z \rangle$ es decreciente (ver la desigualdad (5.17)), y además se encuentra acotada inferiormente ⁴, entonces la sucesión $\{\langle F(x^k), z \rangle\}$ es convergente. Así que

$$\langle F(\hat{x}), z \rangle = \lim_{k \rightarrow +\infty} \langle F(x^k), z \rangle := \inf_{k \in \mathbb{N}} \{\langle F(x^k), z \rangle\} \leq \langle F(x^k), z \rangle.$$

De la expresión anterior se tiene que

$$\langle F(x^k) - F(\hat{x}), z \rangle \geq 0 \quad \forall k \in \mathbb{N} \text{ y } \forall z \in \mathbb{R}_{\geq}^p \setminus \{0\},$$

lo cual implica que $F(x^k) - F(\hat{x}) \in \mathbb{R}_{\geq}^p$, es decir $F(\hat{x}) \preceq F(x^k)$ para todo $k \in \mathbb{N}$, así que por (5.16), $\hat{x} \in E$. \square

Proposición 5.2.1. Sean $\{x^k\}$ y $\{e^k\}$ las sucesiones generadas por el MPPEI. Si las hipótesis **(H1)**, **(H2)**, **(H3)** y **(H4)** son satisfechas, entonces para todo $x \in E$ y para todo $k \in \mathbb{Z}_0^+$ se tiene que

$$\|x^{k+1} - x\|^2 \leq \|x^k - x\|^2 - \|x^{k+1} - x^k\|^2 - \frac{2}{\alpha_k} \langle e^{k+1}, x - x^{k+1} \rangle. \quad (5.18)$$

Demostración. Por la relación (5.9) y el Teorema 5.2.1, en el cuál se demostró la existencia de la sucesión $\{x^k\}$ generada por el algoritmo MPPEI, se tiene que

$$e^{k+1} \in \partial^0 \left(\langle F(\cdot), z_k \rangle + \frac{\alpha_k}{2} \|\cdot - x^k\|^2 \right) (x^{k+1}) + \mathcal{N}_{\Omega_k}(x^{k+1}).$$

Dado que $\langle F(\cdot), z_k \rangle$ es una función localmente Lipschitz (ver las relaciones (5.13) y (5.14)) y $\|\cdot - x^k\|^2$ es convexa diferenciable, por el Lema 3.1.4 se tiene que

$$\begin{aligned} e^{k+1} &\in \partial^0(\langle F(\cdot), z_k \rangle)(x^{k+1}) + \frac{\alpha_k}{2} \nabla \|x^{k+1} - x^k\|^2 + N_{\Omega_k}(x^{k+1}) = \\ &\partial^0 \left(\sum_{i=1}^p z_k^i F_i(\cdot) \right) (x^{k+1}) + \frac{\alpha_k}{2} 2(x^{k+1} - x^k) + N_{\Omega_k}(x^{k+1}). \end{aligned} \quad (5.19)$$

Aplicando respectivamente los Lemas 3.1.3 y 3.1.2, a uno de los términos de (5.19), tenemos

$$\partial^0 \left(\sum_{i=1}^p z_k^i F_i(\cdot) \right) (x^{k+1}) \subset \sum_{i=1}^p \partial^0(z_k^i F_i(x^{k+1})) = \sum_{i=1}^p z_k^i \partial^0 F_i(x^{k+1}),$$

así que en (5.19)

$$e^{k+1} \in \sum_{i=1}^p z_k^i \partial^0 F_i(x^{k+1}) + \alpha_k(x^{k+1} - x^k) + N_{\Omega_k}(x^{k+1}).$$

³Observar que $\Omega_k = \{x \in \mathbb{R}^n : F(x) \preceq F(x^k)\}$.

⁴Dado que $0 \preceq F$ (ver **(H1)**) y $z \in \mathbb{R}_{\geq}^p \setminus \{0\}$, entonces $0 \leq \langle F(\cdot), z \rangle$.

La expresión anterior indica que existen vectores $g_i^k \in \partial^0 F_i(x^{k+1})$ y $v_k \in N_{\Omega_k}(x^{k+1})$, tales que

$$e^{k+1} = \sum_{i=1}^p z_k^i g_i^k + \alpha_k(x^{k+1} - x^k) + v_k,$$

de donde

$$x^k - x^{k+1} = \frac{1}{\alpha_k} \left(\sum_{i=1}^p z_k^i g_i^k + v_k - e^{k+1} \right). \quad (5.20)$$

Ahora tomemos un punto arbitrario $\bar{x} \in E$. Por definición de E (ver (5.16)) se tiene que $\bar{x} \in \Omega_k$ para todo k , en particular $\bar{x} \in \Omega_{k+1}$. Como E es un conjunto cerrado no vacío, $\text{int } \Omega_{k+1} \neq \emptyset$, así que existe $\{x^l\} \subset \text{int } \Omega_{k+1}$ tal que $x^l \rightarrow \bar{x}$. Por otro lado es claro que

$$\|x^k - x^l\|^2 = \|x^k - x^{k+1}\|^2 + \|x^{k+1} - x^l\|^2 + 2\langle x^k - x^{k+1}, x^{k+1} - x^l \rangle.$$

Reemplazando (5.20) en la expresión anterior y haciendo que $x = x^l$, se tiene que

$$\begin{aligned} \|x^k - x^l\|^2 &= \|x^k - x^{k+1}\|^2 + \|x^{k+1} - x^l\|^2 + \\ &\quad \frac{2}{\alpha_k} \left(\sum_{i=1}^p z_k^i \langle g_i^k, x^{k+1} - x^l \rangle + \langle v_k, x^{k+1} - x^l \rangle - \langle e^{k+1}, x^{k+1} - x^l \rangle \right). \end{aligned} \quad (5.21)$$

Dado que F es una función propia, cuasi-convexa y localmente Lipschitz, y además $g_i^k \in \partial^0 F_i(x^{k+1})$ y $F_i(x^l) < F_i(x^{k+1})$ ⁵ para todo $i \in \{1, \dots, p\}$, por la Proposición 3.1.6 se tiene que

$$\begin{aligned} \langle g_i^k, x^l - x^{k+1} \rangle &\leq 0 \\ \langle g_i^k, x^{k+1} - x^l \rangle &\geq 0 \quad \forall i \in \{1, \dots, p\}. \end{aligned} \quad (5.22)$$

Por otro lado, dado que $v_k \in N_{\Omega_k}(x^{k+1})$, entonces $\langle v_k, w - x^{k+1} \rangle \leq 0$ para todo $w \in \Omega_k \subseteq \Omega_{k+1}$ (ver Definición 2.2.7), en particular para $x^l \in \Omega_{k+1}$; entonces

$$\begin{aligned} \langle v_k, x^l - x^{k+1} \rangle &\leq 0 \\ \langle v_k, x^{k+1} - x^l \rangle &\geq 0 \quad \forall i \in \{1, \dots, p\}. \end{aligned} \quad (5.23)$$

Reemplazando las desigualdades (5.23) y (5.22) en (5.21), se tiene que

$$\|x^k - x^l\|^2 \geq \|x^k - x^{k+1}\|^2 + \|x^{k+1} - x^l\|^2 - \frac{2}{\alpha_k} \langle e^{k+1}, x^{k+1} - x^l \rangle \quad \forall k \in \mathbb{N}.$$

Seguidamente, haciendo que $l \rightarrow +\infty$ (y considerando el hecho que $x^l \rightarrow \bar{x}$), en la expresión anterior se tiene que

$$\|x^{k+1} - \bar{x}\|^2 \leq \|x^k - \bar{x}\|^2 - \|x^{k+1} - x^k\|^2 - \frac{2}{\alpha_k} \langle e^{k+1}, \bar{x} - x^{k+1} \rangle \quad \forall k \in \mathbb{N},$$

de donde claramente se sigue el resultado, dado que $\bar{x} \in E$ representa un punto arbitrario de dicho conjunto. \square

⁵Como $x^l \in \text{int } \Omega_{k+1}$ entonces $F(x^l) < F(x^{k+1})$, así que $F_i(x^l) < F_i(x^{k+1})$ para todo $i \in \{1, \dots, p\}$.

5.3. Variantes inexactas del algoritmo MPPEI para minimización multiobjetivo

En esta sección siguiendo un enfoque semejante al realizado por Baygorrea et al. (2016a) ⁶, presentamos y analizamos los resultados de convergencia de dos variantes del algoritmo MPPEI (obtenidas al agregar al mismo dos criterios de error), el primero de los cuáles es una versión clásica introducida por Rockafellar (1976), y el segundo de ellos, estudiado por Tang y Huang (2013). También, bajo ciertas suposiciones (como por ejemplo la acotación del parámetro α_k), probamos que cualquier punto de acumulación de la sucesión generada por cada una de estas variantes del MPPEI, es un punto crítico Pareto-Clarke del POM (5.7).

5.3.1. Algoritmo MPPEI1 y resultados de convergencia

En esta subsección discutimos el algoritmo MPPEI considerando la versión clásica del criterio de estimación de error, a lo que llamaremos el algoritmo MPPEI1, definido por:

Algoritmo MPPEI1

La sucesión $\{x^k\}$ generada por el algoritmo MPPEI (ver las relaciones (5.8) y (5.9)), para el término del error e^{k+1} satisface la siguiente hipótesis:

$$(E1) \quad \sum_{k=0}^{+\infty} \frac{\|e^{k+1}\|}{\alpha_k} < +\infty$$

Observación 5.3.1. De acuerdo con la hipótesis (E1), es importante observar que existe un cierto margen para elegir e^{k+1} . De hecho, esta suposición fue considerada natural en el algoritmo de punto proximal inexacto clásico introducido por Rockafellar (1976).

Lema 5.3.1. Sean $\{x^k\}$ y $\{e^k\}$ las sucesiones generadas por el MPPEI1. Si las hipótesis (H1), (H2), (H3) y (H4) son satisfechas, entonces:

- i) La sucesión $\{x^k\}$ es acotada.
- ii) $\lim_{k \rightarrow +\infty} \|x^{k+1} - x^k\| = 0$.
- iii) Si $\lim_{j \rightarrow +\infty} x^{k_j} = \hat{x}$, entonces $\lim_{j \rightarrow +\infty} x^{k_j+1} = \hat{x}$.

Demostración. i) Sea $\bar{x} \in E$ un punto fijo. Por la Proposición 5.2.1 se tiene que

$$\|x^{k+1} - \bar{x}\|^2 \leq \|x^k - \bar{x}\|^2 - \|x^{k+1} - x^k\|^2 - \frac{2}{\alpha_k} \langle e^{k+1}, \bar{x} - x^{k+1} \rangle,$$

⁶Dichos autores trabajaron con funciones mono-objetivo definidas en espacios de Hadamard.

de donde

$$0 \leq \|x^{k+1} - x^k\|^2 \leq \|x^k - \bar{x}\|^2 - \|x^{k+1} - \bar{x}\|^2 - \frac{2}{\alpha_k} \langle e^{k+1}, \bar{x} - x^{k+1} \rangle, \quad (5.24)$$

así que

$$\|x^{k+1} - \bar{x}\|^2 \leq \|x^k - \bar{x}\|^2 + \frac{2}{\alpha_k} \langle e^{k+1}, x^{k+1} - \bar{x} \rangle.$$

Aplicando la desigualdad de Cauchy-Schwarz (ver Teorema 2.1.2), tenemos

$$\|x^{k+1} - \bar{x}\|^2 \leq \|x^k - \bar{x}\|^2 + 2 \frac{\|e^{k+1}\|}{\alpha_k} \|x^{k+1} - \bar{x}\|. \quad (5.25)$$

Dado que $(a - \frac{1}{2})^2 = a^2 - a + \frac{1}{4} \geq 0$ para todo $a \in \mathbb{R}$, entonces $a \leq a^2 + \frac{1}{4}$. Así que si $a = \|x^{k+1} - \bar{x}\|$,

$$\|x^{k+1} - \bar{x}\| \leq \|x^{k+1} - \bar{x}\|^2 + \frac{1}{4}.$$

Reemplazando la expresión anterior en (5.25), se tiene que

$$\begin{aligned} \|x^{k+1} - \bar{x}\|^2 &\leq \|x^k - \bar{x}\|^2 + 2 \frac{\|e^{k+1}\|}{\alpha_k} \|x^{k+1} - \bar{x}\|^2 + \frac{2}{4} \frac{\|e^{k+1}\|}{\alpha_k} \\ \left(1 - 2 \frac{\|e^{k+1}\|}{\alpha_k}\right) \|x^{k+1} - \bar{x}\|^2 &\leq \|x^k - \bar{x}\|^2 + \frac{1}{2} \frac{\|e^{k+1}\|}{\alpha_k}. \end{aligned}$$

Haciendo que $E^k = \frac{\|e^{k+1}\|}{\alpha_k}$, tenemos

$$\|x^{k+1} - \bar{x}\|^2 \leq \left(1 + \frac{2E^k}{1 - 2E^k}\right) \|x^k - \bar{x}\|^2 + \frac{E^k}{2} \left(\frac{1}{1 - 2E^k}\right).$$

Como $E^k \rightarrow 0$, entonces existe $k_0 \in \mathbb{N}$, tal que $E^k < 1/4$ para todo $k \geq k_0$, lo cual implica que $\left(\frac{1}{1 - 2E^k}\right) < 2$, y que $\left(1 + \frac{2E^k}{1 - 2E^k}\right) < 1 + 2(2E^k)$. Reemplazando estas últimas desigualdades en la expresión anterior, tenemos que

$$\|x^{k+1} - \bar{x}\|^2 \leq^7 (1 + 4E^k) \|x^k - \bar{x}\|^2 + E^k.$$

Finalmente haciendo uso del Lema 2.1.1 se tiene que $\|x^{k+1} - \bar{x}\|^2$ converge, dado que en la expresión anterior los términos $v_k = \|x^{k+1} - \bar{x}\|^2$, $v_{k-1} = \|x^k - \bar{x}\|^2$, $\lambda_k = 4E^k$ y $\beta_k = E^k$ satisfacen lo siguiente: $\sum_{k=1}^{+\infty} \lambda_k < \infty$ y $\sum_{k=1}^{+\infty} \beta_k < \infty$. Por lo tanto la sucesión $\{x^k\}$ es acotada.

ii) Usando nuevamente la Proposición 5.2.1, se tiene que

$$\|x^{k+1} - x^k\|^2 \leq \|x^k - \bar{x}\|^2 - \|x^{k+1} - \bar{x}\|^2 + \frac{2}{\alpha_k} \langle e^{k+1}, x^{k+1} - \bar{x} \rangle.$$

Por la desigualdad de Cauchy-Schwarz (ver Teorema 2.1.2), tenemos

$$\|x^{k+1} - x^k\|^2 \leq \|x^k - \bar{x}\|^2 - \|x^{k+1} - \bar{x}\|^2 + 2 \frac{\|e^{k+1}\|}{\alpha_k} \|x^{k+1} - \bar{x}\|.$$

⁷Ojo esta desigualdad no es estricta, como mostramos en la siguiente situación: Sean $a, b \in \mathbb{R}$ tal que $a < b$, entonces $a \leq b\|\cdot\|$ para toda norma $\|\cdot\|$, dado que $\|\cdot\| \geq 0$.

Aplicando sumatorias, y teniendo en cuenta que

$$\begin{aligned} \sum_{k=0}^m \left(\frac{\|e^{k+1}\|}{\alpha_k} \|x^{k+1} - \bar{x}\| \right) &\leq \sum_{k=0}^m \left(\frac{\|e^{k+1}\|}{\alpha_k} \max_{0 \leq k \leq m} \{\|x^{k+1} - \bar{x}\|\} \right) \\ &= \max_{0 \leq k \leq m} \{\|x^{k+1} - \bar{x}\|\} \sum_{k=0}^m \frac{\|e^{k+1}\|}{\alpha_k}, \end{aligned} \quad (5.26)$$

en la expresión anterior tenemos

$$\sum_{k=0}^m \|x^{k+1} - x^k\|^2 \leq \|x^0 - \bar{x}\|^2 - \|x^{m+1} - \bar{x}\|^2 + 2 \max_{0 \leq k \leq m} \{\|x^{k+1} - \bar{x}\|\} \sum_{k=0}^m \frac{\|e^{k+1}\|}{\alpha_k}.$$

Haciendo que $m \rightarrow \infty$, y teniendo en cuenta la hipótesis **(E1)** y el resultado anterior (ítem (i)) se tiene que $\sum_{k=0}^{\infty} \|x^{k+1} - x^k\|^2 < +\infty$, de donde se sigue el resultado.

iii) Asumamos que $\lim_{k \rightarrow +\infty} x^{k_j} = \hat{x}$. Por el Teorema 2.1.3 (3), es claro que

$$\|x^{k_j+1} - \hat{x}\| \leq \|x^{k_j+1} - x^{k_j}\| + \|x^{k_j} - \hat{x}\|.$$

Tomando límite cuando $j \rightarrow +\infty$, y considerando el ítem (ii) y la hipótesis $\lim_{k \rightarrow +\infty} x^{k_j} = \hat{x}$, en la expresión anterior se tiene que $\lim_{j \rightarrow +\infty} \|x^{k_j+1} - \hat{x}\| \leq 0$, de donde se sigue el resultado. \square

Proposición 5.3.1. Sean $\{x^k\}$ y $\{e^k\}$ las sucesiones generadas por el MPPEI1. Si las hipótesis **(H1)**, **(H2)**, **(H3)** y **(H4)** son satisfechas, entonces la sucesión $\{x^k\}$ converge a un punto de E .

Demostración. Por el Lema 5.3.1 (i) se tiene que la sucesión $\{x^k\}$ es acotada, entonces existe una subsucesión $\{x^{k_j}\} \subseteq \{x^k\}$ y al menos un punto de acumulación $\hat{x} \in \mathbb{R}^n$, tales que $x^{k_j} \rightarrow \hat{x}$ (ver Teorema 2.1.4). Así que por el Lema 5.2.1 se tiene que $\hat{x} \in E$, lo cual implica el resultado. \square

A continuación probaremos que las sucesiones iterantes convergen a un punto crítico Pareto-Clarke, cuando la sucesión de parámetros regulares $\{\alpha_k\}$ es acotada.

Teorema 5.3.1. Si la función $F : \mathbb{R}^n \rightarrow \mathbb{R}^p$, tal que $F_i : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, satisface las hipótesis **(H1)**, **(H2)**, **(H3)** y **(H4)**, con $0 < \alpha_k \leq \bar{\alpha}$, entonces la sucesión $\{x^k\}$ generada por MPPEI1 converge hacia algún punto crítico Pareto-Clarke del POM (5.7).

Demostración. Por la Proposición 5.3.1 existe $\hat{x} \in E$ tal que $x^k \rightarrow \hat{x}$. Probaremos que \hat{x} es un punto crítico Pareto-Clarke (ver Definición 3.2.1). Para llegar a una contradicción asumimos lo contrario, entonces existe algún vector $d \in \mathbb{R}^n$, tal que

$$F_i^0(\hat{x}, d) < 0 \quad \forall i \in \{1, \dots, p\}, \quad (5.27)$$

lo cual implica que d es una dirección de descenso para la función multiobjetivo F en \hat{x} , así que existe $\delta > 0$ tal que

$$F(\hat{x} + \lambda d) < F(\hat{x}) \quad \forall \lambda \in (0, \delta].$$

Por lo tanto $(\hat{x} + \lambda d) \in \Omega_k$.

Por otro lado, recordemos que por el Teorema 5.2.1 la sucesión $\{x^{k+1}\}$ generada por el MPPEI existe, y además $x^{k+1} \in \Omega_k$. En la Proposición 5.2.1 - relación (5.19) también se probó que

$$e^{k+1} \in \partial^0(\langle F(\cdot), z_k \rangle)(x^{k+1}) + \alpha_k(x^{k+1} - x^k) + N_{\Omega_k}(x^{k+1}),$$

así que

$$\alpha_k(x^k - x^{k+1}) - v_k + e^{k+1} \in \partial^0(\langle F(\cdot), z_k \rangle)(x^{k+1}),$$

donde $v_k \in N_{\Omega_k}(x^{k+1})$. Haciendo uso de la Definición 3.1.2(2) concerniente al término $\partial^0(\cdot)$, en la expresión anterior se tiene que

$$\langle \alpha_k(x^k - x^{k+1}) - v_k + e^{k+1}, q \rangle \leq \langle F(\cdot), z_k \rangle^0(x^{k+1}, q) \quad \forall q \in \mathbb{R}^n, \quad (5.28)$$

donde $\langle F(\cdot), z_k \rangle^0(x^{k+1}, q)$ representa la derivada direccional de Clarke. Utilizando propiedades básicas del producto interno de vectores, y teniendo en cuenta que $\alpha_k > 0$, en (5.28) tenemos

$$\alpha_k \langle x^k - x^{k+1}, q \rangle - \langle v_k, q \rangle + \langle e^{k+1}, q \rangle \leq \langle F(\cdot), z_k \rangle^0(x^{k+1}, q) \quad \forall q \in \mathbb{R}^n. \quad (5.29)$$

Considerando que $q = (\hat{x} + \lambda d - x^{k+1})^8$, y teniendo en cuenta que $v_k \in N_{\Omega_k}(x^{k+1})^9$, se tiene que $\langle v_k, q \rangle \leq 0$ (o equivalente $0 \leq -\langle v_k, q \rangle$), lo cual se justifica por el hecho de que $(\hat{x} + \lambda d) \in \Omega_k$. Así en (5.29)

$$\alpha_k \langle x^k - x^{k+1}, q \rangle + \langle e^{k+1}, q \rangle \leq \langle F(\cdot), z_k \rangle^0(x^{k+1}, q). \quad (5.30)$$

Aplicando respectivamente los ítems (i), (ii) y (iii) del Lema 3.1.1, en el término de la derecha de la desigualdad anterior, se tiene

$$\begin{aligned} (\langle F(\cdot), z_k \rangle)^0(x^{k+1}, q) &= \left(\sum_{i=1}^p z_k^i F_i(\cdot) \right)^0(x^{k+1}, q) \stackrel{(i)}{\leq} \sum_{i=1}^p (z_k^i F_i)^0(x^{k+1}, q) \\ &\stackrel{(ii)}{=} \sum_{i=1}^p z_k^i F_i^0(x^{k+1}, q) \stackrel{(iii)}{=} \sum_{i=1}^p F_i^0(x^{k+1}, z_k^i q), \end{aligned}$$

donde z_k^i son las componentes del vector z_k . Reemplazando la expresión anterior en (5.30), se tiene que

$$\langle \alpha_k(x^k - x^{k+1}) + e^{k+1}, q \rangle = \alpha_k \langle x^k - x^{k+1}, q \rangle + \langle e^{k+1}, q \rangle \leq \sum_{i=1}^p F_i^0(x^{k+1}, z_k^i q). \quad (5.31)$$

⁸Esta construcción se debe al hecho que $(\hat{x} + \lambda d) \in \Omega_k$ y $x^{k+1} \in \Omega_k$.

⁹ $N_{\Omega_k}(x^{k+1}) = \{d \in \mathbb{R}^n : \langle d, x - x^{k+1} \rangle \leq 0, \forall x \in \Omega_k\}$ - ver Definición 2.2.7.

Por la desigualdad de Cauchy Schwarz (ver Teorema 2.1.2) y algunas propiedades básicas de la norma de vectores, en el término de la izquierda de (5.31) se tiene que

$$\begin{aligned} |\langle \alpha_k(x^k - x^{k+1}) + e^{k+1}, q \rangle| &\leq \|\alpha_k(x^k - x^{k+1}) + e^{k+1}\| \|q\| \\ &\leq \alpha_k \|x^k - x^{k+1}\| \|q\| + \|e^{k+1}\| \|q\|. \end{aligned}$$

Aplicando al término de la izquierda de la expresión anterior la siguiente propiedad: para todo $a, b \in \mathbb{R}$, $|x| \leq a \iff [(a \geq 0) \wedge (-a \leq x \leq a)]$, y reemplazando dicho resultado en (5.31), tenemos

$$-\alpha_k \|x^k - x^{k+1}\| \|q\| - \alpha_k \left(\frac{\|e^{k+1}\|}{\alpha_k} \right) \|q\| \leq \sum_{i=1}^p F_i^0(x^{k+1}, z_{k_j}^i q). \quad (5.32)$$

Como $\{z_k\}$ es acotada, por el Teorema (2.1.4) existe una subsucesión $\{z_{k_j}\}$ tal que $\lim_{j \rightarrow +\infty} z_{k_j} = \bar{z} \in \mathbb{R}_{\geq}^p \setminus \{0\}$. De ser necesario, y sin pérdida de generalidad, también podemos considerar las subsucesiones $\{x^{k_j}\} \subseteq \{x^k\}$ y $\{\alpha_{k_j}\} \subseteq \{\alpha_k\}$. Además, como $0 < \alpha_k \leq \bar{\alpha}$, en (5.32) se tiene que

$$-\bar{\alpha} \|x^{k_j} - x^{k_j+1}\| \|q'\| - \bar{\alpha} \left(\frac{\|e^{k_j+1}\|}{\alpha_{k_j}} \right) \|q'\| \leq \sum_{i=1}^p F_i^0(x^{k_j+1}, z_{k_j}^i q'), \quad (5.33)$$

donde $q' = (\hat{x} + \lambda d - x^{k_j+1})$ y $z_{k_j}^i$ son las componentes del vector z_{k_j} . Tomando límite superior cuando $j \rightarrow +\infty$, se puede verificar que el término de la izquierda de la desigualdad anterior tiende a cero. Esto se da en base a los ítems (ii) y (iii) del Lema 5.3.1¹⁰, a la hipótesis **(E1)**, y al hecho que $\lambda > 0$ y $z_{k_j} \rightarrow \bar{z}$. Con las mismas condiciones, a continuación analizamos el término de la derecha de (5.33):

$$0 \leq \limsup_{j \rightarrow +\infty} \sum_{i=1}^p F_i^0(x^{k_j+1}, z_{k_j}^i q') = \sum_{i=1}^p \limsup_{j \rightarrow +\infty} F_i^0(x^{k_j+1}, z_{k_j}^i q'). \quad (5.34)$$

Dado que F es propia y localmente Lipschitz, por la Proposición 3.1.5 se tiene que F_i^0 es semicontinua superior. Entonces

$$\limsup_{j \rightarrow +\infty} F_i^0(x^{k_j+1}, z_{k_j}^i q') \leq F_i^0(\hat{x}, \bar{z}^i \lambda d),$$

donde $\{(x^{k_j+1}, z_{k_j}^i q')\}$ es una sucesión en $\mathbb{R}^n \times \mathbb{R}^n$, la cuál por las condiciones anteriormente tratadas satisface $\lim_{j \rightarrow +\infty} (x^{k_j+1}, z_{k_j}^i q') = (\hat{x}, \bar{z}^i \lambda d)$. Reemplazando la desigualdad anterior en (5.34), tenemos

$$0 \leq \sum_{i=1}^p \limsup_{j \rightarrow +\infty} F_i^0(x^{k_j+1}, z_{k_j}^i q') \leq \sum_{i=1}^p F_i^0(\hat{x}, \bar{z}^i \lambda d). \quad (5.35)$$

Haciendo uso del Lema 3.1.1 (iii), y teniendo en cuenta que $\lambda > 0$, en (5.35) se tiene que

$$\sum_{i=1}^p \bar{z}^i F_i^0(\hat{x}, d) = \bar{z}^1 F_1^0(\hat{x}, d) + \cdots + \bar{z}^p F_p^0(\hat{x}, d) \geq 0. \quad (5.36)$$

¹⁰De (ii) se tiene que $\lim_{j \rightarrow +\infty} \|x^{k_j+1} - x^{k_j}\| = 0$, y de (iii) $\lim_{j \rightarrow +\infty} x^{k_j+1} = \hat{x}$, así que $q' = (\hat{x} + \lambda d - x^{k_j+1}) \rightarrow \lambda d$, donde $\lambda \in (0, \delta]$.

Como $\bar{z} \in \mathbb{R}_{\geq}^p \setminus \{0\}$, entonces existe un conjunto de índices en los cuáles algunas de las componentes de \bar{z} son estrictamente positivas. Sin pérdida de generalidad consideramos $I = \{i \in \{1, \dots, p\} : \bar{z}^i > 0\}$. Entonces en (5.36) existe un $i_0 \in I$ tal que $F_{i_0}^0(\hat{x}, d) \geq 0$, lo cual contradice (5.27). Por lo tanto \hat{x} es un punto crítico Pareto-Clarke del POM (5.7). \square

Corolario 5.3.1. *Sea $F : \mathbb{R}^n \rightarrow \mathbb{R}^p$ una función \mathbb{R}_{\geq}^p -convexa que satisface las hipótesis (H1), (H2), (H3) y (H4), con $0 < \alpha_k \leq \bar{\alpha}$. Si \hat{x} es un punto de convergencia de la sucesión $\{x^k\}$ generada por algoritmo MPPEI1, entonces \hat{x} es una solución débil de Pareto del POM (5.7).*

Demostración. Como \hat{x} es un punto de convergencia de $\{x^k\}$, por el Teorema 5.3.1 \hat{x} es un punto crítico Pareto-Clarke. Finalmente haciendo uso de la Proposición 3.2.1 se tiene que \hat{x} es una solución débil de Pareto del POM (5.7). \square

Corolario 5.3.2. *Si $F : \mathbb{R}^n \rightarrow \mathbb{R}^p$ es una función continuamente diferenciable en \mathbb{R}^n , y además satisface las hipótesis (H1), (H3) y (H4), con $0 < \alpha_k \leq \bar{\alpha}$, entonces el punto de convergencia \hat{x} de la sucesión $\{x^k\}$ generada por el MPPEI1, es un punto crítico de Pareto-Clarke del problema (5.7), es decir, existe $i_0 \in \{1, 2, \dots, p\}$ tal que*

$$\langle \nabla F_{i_0}(\hat{x}), d \rangle \geq 0 \quad \forall d \in \mathbb{R}^n. \quad (5.37)$$

Demostración. Es inmediata, dado que continuamente diferenciable en \mathbb{R}^n implica la suposición (H2). Así que por el Teorema 3.1.1 y la relación (2.3) tenemos que $F_i^0(\hat{x}, d) = F_i'(\hat{x}, d) = \langle \nabla F_i(\hat{x}), d \rangle$, donde F_i' es la derivada direccional de F_i . Finalmente aplicando el Teorema 5.3.1 obtenemos el resultado deseado. \square

5.3.2. Algoritmo MPPEI2 y convergencia de resultados

En este apartado presentamos el MPPEI dotado de un criterio de error estudiado por Tang y Huang (2013), y generalizado al caso cuasiconvexo por Baygorrea et al. (2016a)¹¹, a lo cuál llamaremos MPPEI2, definido por:

Algoritmo MPPEI2

La sucesión $\{x^k\}$ generada por el algoritmo MPPEI (ver las relaciones (5.8) y (5.9)), para el término del error e^{k+1} satisface las siguientes hipótesis:

$$\begin{aligned} \text{(E2)} \quad & \frac{\|e^{k+1}\|}{\alpha_k} \leq \eta_k \|x^{k+1} - x^k\| \\ \text{(E2.1)} \quad & \sum_{k=0}^{+\infty} \eta_k^2 < +\infty \end{aligned}$$

¹¹Cabe resaltar que en ambos casos, los autores trabajaron con funciones mono-objetivos, definidos en espacios de Hadamard.

Lema 5.3.2. Sean $\{x^k\}$ y $\{e^k\}$ las sucesiones generadas por el MPPEI2. Si las hipótesis **(H1)**, **(H2)**, **(H3)** y **(H4)** son satisfechas, entonces existe un entero $k_0 \geq 0$, tal que para todo $k \geq k_0$ se tiene que

$$i) \quad \|x^{k+1} - x\|^2 \leq \left(1 + \frac{2\eta_k^2}{1 - 2\eta_k^2}\right) \|x^k - x\|^2 - \frac{1}{2} \|x^{k+1} - x^k\|^2 \quad \forall x \in E; \quad (5.38)$$

ii) La sucesión $\{x^k\}$ es acotada;

iii) $\lim_{k \rightarrow +\infty} \|x^{k+1} - x^k\| = 0$;

iv) Si $\lim_{j \rightarrow +\infty} x^{k_j} = x$, entonces $\lim_{j \rightarrow +\infty} x^{k_j+1} = x \in E$.

Demostración. i) Probaremos que se verifica la desigualdad (5.38). Por la propiedad de norma de vectores se tiene que $\|y\| = \sqrt{\langle y, y \rangle} \geq 0$ para todo $y \in \mathbb{R}^n$. Dado que $\eta_k \rightarrow 0^+$, definimos $y = \frac{e^{k+1}}{\sqrt{2\alpha_k\eta_k}} - \sqrt{2\alpha_k\eta_k}(x^{k+1} - x)$. Entonces

$$0 \leq \langle y, y \rangle = \frac{1}{2\alpha_k\eta_k^2} \|e^{k+1}\|^2 - 2\langle e^{k+1}, x^{k+1} - x \rangle + 2\alpha_k\eta_k^2 \|x^{k+1} - x\|^2,$$

de donde

$$-2\langle e^{k+1}, x - x^{k+1} \rangle \leq \frac{1}{2\alpha_k\eta_k^2} \|e^{k+1}\|^2 + 2\alpha_k\eta_k^2 \|x^{k+1} - x\|^2. \quad (5.39)$$

Por otro lado, por la desigualdad (5.18) se tiene que

$$\|x^{k+1} - x\|^2 \leq \|x^k - x\|^2 - \|x^{k+1} - x^k\|^2 - \frac{2}{\alpha_k} \langle e^{k+1}, x - x^{k+1} \rangle \quad \forall k \in \mathbb{Z}_0^+. \quad (5.40)$$

Sustituyendo (5.39) en (5.40) se tiene que

$$\|x^{k+1} - x\|^2 \leq \|x^k - x\|^2 - \|x^{k+1} - x^k\|^2 + \frac{1}{2\alpha_k^2\eta_k^2} \|e^{k+1}\|^2 + 2\eta_k^2 \|x^{k+1} - x\|^2.$$

Haciendo uso de la condición **(E2)** tenemos que

$$\|x^{k+1} - x\|^2 \leq \|x^k - x\|^2 - \|x^{k+1} - x^k\|^2 + \frac{1}{2} \|x^{k+1} - x^k\|^2 + 2\eta_k^2 \|x^{k+1} - x\|^2,$$

de donde

$$(1 - 2\eta_k^2) \|x^{k+1} - x\|^2 \leq \|x^k - x\|^2 - \frac{1}{2} \|x^{k+1} - x^k\|^2.$$

Como $\eta_k \rightarrow 0^+$ (esto se cumple por la condición **(E2.1)**), entonces existe $k_0 \geq 0$ tal que $1 - 2\eta_k^2 > 0$ para todo $k \geq k_0$. Así en la expresión anterior tenemos

$$\|x^{k+1} - x\|^2 \leq \left(\frac{1}{1 - 2\eta_k^2}\right) \|x^k - x\|^2 - \left(\frac{1}{2(1 - 2\eta_k^2)}\right) \|x^{k+1} - x^k\|^2.$$

Como $0 < 1 - 2\eta_k^2 < 1$ (lo cual se cumple dado que $\eta_k \rightarrow 0^+$), entonces $\frac{-1}{2(1 - 2\eta_k^2)} < \frac{-1}{2}$. Así en la expresión anterior se tiene que

$$\|x^{k+1} - x\|^2 \leq {}^{12} \left(1 + \frac{2\eta_k^2}{1 - 2\eta_k^2}\right) \|x^k - x\|^2 - \frac{1}{2} \|x^{k+1} - x^k\|^2,$$

¹²Esta desigualdad no es estricta, dado que la norma $\|x^{k+1} - x^k\|^2$ que multiplica al término $1/2(1 - 2\eta_k^2)$, es ≥ 0 .

con lo cual obtenemos la primera parte del teorema.

ii) Probaremos que $\{x^k\}$ es una sucesión acotada. De la desigualdad (5.38), es claro que

$$\|x^{k+1} - x\|^2 \leq \left(1 + \frac{2\eta_k^2}{1 - 2\eta_k^2}\right) \|x^k - x\|^2 \quad \forall k \geq k_0. \quad (5.41)$$

Dado que $\eta_k^2 \rightarrow 0^+$, para todo $\epsilon > 0$ existe $\tilde{k}_0 \geq 0$, tal que $\eta_k^2 < \epsilon$ para todo $k \geq \tilde{k}_0$. Entonces

$$0 < 1 - 2\epsilon < 1 - 2\eta_k^2 < 1.$$

Por lo tanto, para todo $\epsilon < 1/2$ (lo cual se deduce del hecho que $0 < 1 - 2\epsilon$), tenemos

$$\frac{2\eta_k^2}{1 - 2\eta_k^2} < \frac{2\eta_k^2}{1 - 2\epsilon} \quad \forall k \geq \tilde{k}_0.$$

Aplicando sumatorias, y teniendo en cuenta la hipótesis (**E2.1**), se tiene que

$$\sum_{k=1}^{+\infty} \frac{2\eta_k^2}{1 - 2\eta_k^2} < +\infty. \quad (5.42)$$

Finalmente haciendo uso del Lema 2.1.1 se tiene que $\|x^{k+1} - x\|^2$ converge, dado que en la expresión (5.41) los términos $v_k = \|x^{k+1} - x\|^2$, $v_{k-1} = \|x^k - x\|^2$, $\lambda_k = \frac{2\eta_k^2}{1 - 2\eta_k^2}$ y $\beta_k = 0$, satisfacen $\sum_{k=1}^{+\infty} \lambda_k < \infty$ (ver la relación (5.42)) y $\sum_{k=1}^{+\infty} \beta_k < \infty$. Por lo tanto la sucesión $\{x^k\}$ es acotada.

iii) Probaremos que $\lim_{k \rightarrow +\infty} \|x^{k+1} - x^k\| = 0$. De la relación (5.38) se tiene que

$$\frac{1}{2} \|x^{k+1} - x^k\| \leq \|x^k - x\|^2 - \|x^{k+1} - x\|^2 + \frac{2\eta_k^2}{1 - 2\eta_k^2} \|x^k - x\|^2.$$

Aplicando sumatorias, y teniendo en cuenta lo indicado en la relación (5.26), se tiene que

$$\begin{aligned} \frac{1}{2} \sum_{k=k_0}^m \|x^{k+1} - x^k\| &\leq \sum_{k=k_0}^m (\|x^k - x\|^2 - \|x^{k+1} - x\|^2) + \sum_{k=k_0}^m \frac{2\eta_k^2}{1 - 2\eta_k^2} \|x^k - x\|^2 \leq \\ &\|x^{k_0} - x\|^2 - \|x^{m+1} - x\|^2 + \max_{k_0 \leq k \leq m} \{\|x^k - x\|^2\} \sum_{k=k_0}^m \frac{2\eta_k^2}{1 - 2\eta_k^2}. \end{aligned}$$

Dado que $\{x^k\}$ es una sucesión acotada, entonces $\|x^k - x\|$ para tado $k \in \mathbb{N}$ también es acotada. Por lo tanto tomando límite cuando $m \rightarrow +\infty$, y considerando la expresión (5.42), tenemos que $\sum_{k=k_0}^{+\infty} \|x^{k+1} - x^k\| < +\infty$, así que $\lim_{k \rightarrow +\infty} \|x^{k+1} - x^k\| = 0$.

vi) La demostración pendiente es semejante a la prueba realizada en el ítem iii) del Lema 5.3.1. \square

Proposición 5.3.2. Sean $\{x^k\}$ y $\{e^k\}$ las sucesiones generadas por el algoritmo MPPEI2. Si las hipótesis (**H1**), (**H2**), (**H3**) y (**H4**) son satisfechas, entonces la sucesión $\{x^k\}$ converge a un punto de E .

Demostración. Por el Lema 5.3.2 se tiene que la sucesión $\{x^k\}$ es acotada, entonces existe al menos un punto de acumulación $\hat{x} \in \mathbb{R}^n$, así como también una subsucesión $\{x^{k_j}\} \subseteq \{x^k\}$, tales que $x^{k_j} \rightarrow \hat{x}$ (ver Teorema 2.1.4). Entonces por el Lema 5.2.1 se tiene que $\hat{x} \in E$, lo cual implica el resultado. \square

Teorema 5.3.2. *Si la función $F : \mathbb{R}^n \rightarrow \mathbb{R}^p$ satisface las hipótesis (H1), (H2), (H3) y (H4), con $0 < \alpha_k \leq \bar{\alpha}$, entonces la sucesión $\{x^k\}$ generada por el MPPEI2 converge hacia algún punto crítico Pareto-Clarke del POM (5.7).*

Demostración. Considerando que $\{x^k\}$ es la sucesión generada por el algoritmo MPPEI2, tomamos un $\hat{x} \in \mathbb{R}^n$ tal que $x^k \rightarrow \hat{x}$. Para probar que \hat{x} es un punto crítico Pareto Clarke, suponemos lo contrario; entonces siguiendo los mismos pasos realizados en la primera parte de la prueba del Teorema 5.3.1, hasta la relación (5.32), se tiene que

$$-\alpha_k \|x^k - x^{k+1}\| \|q\| - \|e^{k+1}\| \|q\| \leq \sum_{i=1}^p F_i^0(x^{k+1}, z_k^i q).$$

Aplicando la hipótesis (E2), y teniendo en cuenta que $0 < \alpha_k \leq \bar{\alpha}$, en la expresión anterior se tiene que

$$\begin{aligned} -\alpha_k \|x^k - x^{k+1}\| \|q\| - \alpha_k \eta_k \|x^{k+1} - x^k\| \|q\| &\leq \sum_{i=1}^p F_i^0(x^{k+1}, z_k^i q) \\ -\bar{\alpha}(1 + M) \|x^{k+1} - x^k\| \|q\| &\leq \sum_{i=1}^p F_i^0(x^{k+1}, z_k^i q), \end{aligned}$$

para algún $M > 0$ (ojo, por la hipótesis (E2.1) $\eta_k \rightarrow 0^+$). Siguiendo nuevamente los pasos que se encuentran entre las relaciones (5.32) y (5.33) del Teorema 5.3.1, tenemos que

$$-\bar{\alpha}(1 + M) \|x^{k_j+1} - x^{k_j}\| \|q'\| \leq \sum_{i=1}^p F_i^0(x^{k_j+1}, z_{k_j}^i q'),$$

donde $q' = (\hat{x} + \lambda d - x^{k_j+1})$. Tomando límite superior cuando $j \rightarrow +\infty$, y siguiendo lo indicado entre las relaciones (5.33) y (5.36) (considerando en este caso el Lema 5.3.2, en lugar del Lema 5.3.1), se tiene que

$$0 \leq \sum_{i=1}^p \bar{z}^i F_i^0(\hat{x}, d) = \bar{z}^1 F_1^0(\hat{x}, d) + \cdots + \bar{z}^p F_p^0(\hat{x}, d), \quad (5.43)$$

lo cual representa la misma expresión obtenida en la relación (5.36). Finalmente siguiendo lo indicado en el párrafo posterior a (5.36), se tiene que \hat{x} es un punto crítico Pareto-Clarke del POM (5.7). \square

Observación 5.3.2. *Siguiendo procedimientos semejantes a los desarrollados en los Corolarios 5.3.1 y 5.3.1 (pero considerando ahora el Teorema 5.3.2, en lugar del Teorema 5.3.1), podemos obtener resultados semejantes a los planteados en dichos corolarios.*

Capítulo 6

TASA DE CONVERGENCIA DEL ALGORITMO MPPEI2

El objetivo de este capítulo, es analizar la tasa de convergencia del algoritmo MPPEI2, dado por:

Algoritmo MPPEI2

Inicio: Escoger un punto inicial arbitrario

$$x^0 \in \mathbb{R}^n. \quad (6.1)$$

Paso iterativo: Para $k = 0, 1, \dots$, y dado un x^k , encontrar $x^{k+1} \in \Omega_k$ y $e^{k+1} \in \mathbb{R}^n$, tales que:

$$e^{k+1} \in \partial^0 \left(\langle F(\cdot), z_k \rangle + \frac{\alpha_k}{2} \|\cdot - x^k\|^2 \right) (x^{k+1}) + \mathcal{N}_{\Omega_k}(x^{k+1}), \quad (6.2)$$

donde α_k , Ω_k , y z_k están definidos en la Sección 5.2 - página 66. Además, el término del error e^{k+1} satisface las siguientes condiciones:

$$\frac{\|e^{k+1}\|}{\alpha_k} \leq \eta_k \|x^{k+1} - x^k\| \quad (6.3)$$

$$\sum_{k=0}^{+\infty} \eta_k^2 < +\infty \quad (6.4)$$

Criterio de parada: Si $x^{k+1} = x^k$ o si x^{k+1} es un punto crítico Pareto-Clarke detener el criterio. En otro caso hacer $k \leftarrow k + 1$ y retornar al paso iterativo.

Éste algoritmo (como ya se vió en la Sección 5.2), tiene como objetivo resolver el siguiente problema de minimización multiobjetivo cuasi-convexo

$$\min_{x \in \mathbb{R}^n} F(x) = \min\{F(x), x \in \mathbb{R}^n\}, \quad (6.5)$$

donde $F : \mathbb{R}^n \rightarrow \mathbb{R}^p$ es una función que satisface las siguientes hipótesis:

(H1) $0 \preceq F$, es decir que $F_i(x) \geq 0$ para todo $i \in \{1, \dots, p\}$ y $x \in \text{dom}(F_i)$,

(H2) F es localmente Lipschitz en \mathbb{R}^n (ver Definición 2.2.3 y 2.2.13),

(H3) F es \mathbb{R}_{\geq}^p -cuasi-convexa (ver Definición 2.2.11),

(H4) $(F(x^0) - \mathbb{R}_{\geq}^p) \cap F(\mathbb{R}^n)$ es \mathbb{R}_{\geq}^p -completo (ver página 69).

Con tal objetivo, y bajo las hipótesis anteriores, de forma análoga a (5.16), denotamos el siguiente conjunto convexo cerrado no vacío

$$E = \{x \in \mathbb{R}^n : F(x) \preceq F(x^k), \quad \forall k \in \mathbb{N}\}. \quad (6.6)$$

También denotamos el conjunto

$$Z = E \cap PC, \quad (6.7)$$

donde PC representa al conjunto de puntos críticos Pareto-Clarke del POM (6.5) (ver Definición 3.2.1), dado por:

$$PC = \{x \in \mathbb{R}^n : F_{i_0}^0(x, d) \geq 0, \quad \forall d \in \mathbb{R}^n \text{ y para algún } i_0 = i_0(d) = 1, \dots, p\}.$$

Observación 6.0.1. De (6.2) y del hecho que $\partial^{Lim}(\delta_{\Omega_k})(x^{k+1}) = N_{\Omega_k}(x^{k+1})$ (ver Proposición 3.1.3), se tiene que

$$e^{k+1} \in \partial^0 \left(\langle F(\cdot), z_k \rangle + \frac{\alpha_k}{2} \|\cdot - x^k\|^2 \right) (x^{k+1}) + \partial^{Lim}(\delta_{\Omega_k})(x^{k+1}).$$

Dado que $\langle F(\cdot), z_k \rangle$ es una función localmente Lipschitz (ver las relaciones (5.13) y (5.14)) y $\|\cdot - x^k\|^2$ es convexa diferenciable. Aplicando el Lema 3.1.4 a la expresión anterior se tiene que

$$\begin{aligned} e^{k+1} &\in \partial^0(\langle F(\cdot), z_k \rangle)(x^{k+1}) + \frac{\alpha_k}{2} \nabla \|x^{k+1} - x^k\|^2 + \partial^{Lim}(\delta_{\Omega_k})(x^{k+1}) = \\ &\partial^0(\langle F(\cdot), z_k \rangle)(x^{k+1}) + \frac{\alpha_k}{2} 2(x^{k+1} - x^k) + \partial^{Lim}(\delta_{\Omega_k})(x^{k+1}), \end{aligned}$$

de donde

$$e^{k+1} + \alpha_k(x^k - x^{k+1}) \in \partial^0(\langle F(\cdot), z_k \rangle)(x^{k+1}) + \partial^{Lim}(\delta_{\Omega_k})(x^{k+1}).$$

Así que existe $g^k \in (\partial^0(\langle F(\cdot), z_k \rangle) + \partial^{Lim}\delta_{\Omega_k}(\cdot))(x^{k+1})$, tal que

$$g^k = e^{k+1} + \alpha_k(x^k - x^{k+1}). \quad (6.8)$$

6.1. Análisis de la tasa de convergencia

Para este propósito vamos a considerar la siguiente hipótesis:

(H5) Para $\bar{x} \in Z$ tal que $x^k \rightarrow \bar{x}$, existen $\delta = \delta(\bar{x}) > 0$ y $\tau = \tau(\bar{x}) > 0$, tales que para todo $w \in B(0, \delta) \subset \mathbb{R}^n$ y para todo x^{k+1} con $w \in (\partial^0(\langle F(\cdot), z_k \rangle) + \partial^{Lim}\delta_{\Omega_k}(\cdot))(x^{k+1})$, se tiene que

$$\|x^{k+1} - \bar{x}\| \leq \tau \|w\|. \quad (6.9)$$

Observación 6.1.1. La hipótesis anterior, a la cuál también la podemos llamar la condición de crecimiento en el punto de convergencia $\bar{x} \in \mathbb{R}^n$, es diferente a la propuesta por Tang y Huang (2013)¹, y a la de Baygorrea (2017) (quien la generalizada al campo de las funciones cuasi-convexas), dado que ambos trabajan con funciones mono-objetivos definidas en espacios de Hadamard. De hecho, una vez que la sucesión $\{x^k\}$ es obtenida, la hipótesis **(H5)** establece una propiedad de crecimiento del operador subdiferencial en el punto $x^k \in \mathbb{R}^n$, lejos del punto solución \bar{x} .

El conjunto Z (ver relación (6.7)) y la hipótesis **(H5)** que se está proponiendo para analizar la tasa de convergencia del algoritmo MPPEI2, representan una generalización hacia el campo multiobjetivo, de los trabajos de Tang y Huang (2013) y de Baygorrea (2017).

Lema 6.1.1. Sean $\{x^k\}$ y $\{e^k\}$ las sucesiones generadas por el MPPEI2. Supongamos que las hipótesis **(H1)**, **(H2)**, **(H3)**, **(H4)** y **(H5)** son satisfechas, con $0 < \alpha_k \leq \bar{\alpha}$, entonces

i) existe $\tilde{k} \in \mathbb{N}$ tal que

$$\|g^k\| < \delta \quad \forall k \geq \tilde{k}, \quad (6.10)$$

donde g^k es dado por (6.8);

ii) se tiene que

$$\|x^{k+1} - \bar{x}\| \leq \tau \alpha_k (\eta_k + 1) \|x^{k+1} - x^k\| \quad \forall k \geq \tilde{k}. \quad (6.11)$$

Demostración. i) Sean $\bar{x} = \lim_{k \rightarrow +\infty} x^k$ y $g^k \in (\partial^0(\langle F(\cdot), z_k \rangle) + \partial^{Lim} \delta_{\Omega_k}(\cdot))(x^{k+1})$ dado por (6.8). Haciendo uso de algunas propiedades básicas de norma de vectores, y considerando la la relación (6.3) y el hecho que $\alpha_k \leq \bar{\alpha}$, se tiene que

$$\begin{aligned} \|g^k\| &\leq \|e^{k+1}\| + \alpha_k \|x^k - x^{k+1}\| \\ &\leq (\alpha_k \eta_k + \alpha_k) \|x^{k+1} - x^k\| \end{aligned} \quad (6.12)$$

$$\leq \bar{\alpha} (\eta_k + 1) \|x^{k+1} - x^k\| \quad k \geq 0. \quad (6.13)$$

Como $\eta_k \rightarrow 0$ y $\|x^{k+1} - x^k\| \rightarrow 0$ (ver Lema 5.3.2), tomando $\epsilon = \delta$, existe un $\tilde{k} \in \mathbb{N}$ tal que $\|g^k\| < \delta$ para todo $k \geq \tilde{k}$.

ii) En (6.9) haciendo que $w = g^k$ para todo $k \geq \tilde{k}$, se tiene que:

$$\|x^{k+1} - \bar{x}\| \leq \tau \|w\| = \tau \|g^k\|. \quad (6.14)$$

Por lo tanto la relación (6.11) es obtenida reemplazando (6.12) en la expresión anterior. \square

¹Lo propuso para trabajar problemas de singularidad de campos vectoriales monótonos maximales, en particularmente para problemas de minimización convexa en variedades de Hadamard.

Ahora vamos a presentar el teorema relacionado a la tasa de convergencia del algoritmo MPPEI2, completando así el resultado de convergencia dado por la Proposición 5.3.2.

Teorema 6.1.1. Sean $\{x^k\}$ y $\{e^k\}$ las sucesiones generadas por el MPPEI2. Supongamos que las hipótesis **(H1)**, **(H2)**, **(H3)**, **(H4)** y **(H5)** son satisfechas, con $0 < \alpha_k \leq \bar{\alpha}$, entonces $\{x^k\}$ converge linealmente hacia $\bar{x} \in Z$. Además de ello, si $\alpha_k \searrow 0$, entonces la convergencia es superlineal.

Demostración. Sean $\bar{x} \in Z$ un punto límite de la sucesión $\{x^k\}$ y $g^k \in (\partial^0(\langle F(\cdot), z_k \rangle) + \partial^{Lim} \delta_{\Omega_k}(\cdot))(x^{k+1})$ dado por (6.8). Definimos

$$w^k = g^k.$$

Debido a la relación (6.10) tenemos que $\|w^k\| = \|g^k\| < \delta$ para todo $k \geq \tilde{k}$; así que $w^k \in B(0, \delta)$ para todo $k \geq \tilde{k}$. Además, $g^k \in (\partial^0(\langle F(\cdot), z_k \rangle) + \partial^{Lim} \delta_{\Omega_k}(\cdot))(x^{k+1})$.

Como $\bar{x} \in Z$, entonces $\bar{x} \in E$ (ver la relación (6.7)); así que por (5.38) tenemos que

$$\|x^{k+1} - \bar{x}\|^2 \leq \left(1 + \frac{2\eta_k^2}{1 - 2\eta_k^2}\right) \|x^k - \bar{x}\|^2 - \frac{1}{2} \|x^{k+1} - x^k\|^2 \quad \forall k \geq k_0. \quad (6.15)$$

Por la relación (6.11) también tenemos

$$\|x^{k+1} - x^k\| \geq \frac{1}{\tau \alpha_k (\eta_k + 1)} \|x^{k+1} - \bar{x}\| \quad \forall k \geq \tilde{k},$$

de donde

$$-\|x^{k+1} - x^k\|^2 \leq -\left(\frac{1}{\tau \alpha_k (\eta_k + 1)}\right)^2 \|x^{k+1} - \bar{x}\|^2 \quad \forall k \geq \tilde{k}.$$

Reemplazando la relación anterior en (6.15), se tiene que

$$\|x^{k+1} - \bar{x}\|^2 \leq \left(\frac{1}{1 - 2\eta_k^2}\right) \|x^k - \bar{x}\|^2 - \frac{1}{2} \left(\frac{1}{\tau \alpha_k (\eta_k + 1)}\right)^2 \|x^{k+1} - \bar{x}\|^2 \quad \forall k \geq \max\{k_0, \tilde{k}\},$$

así que

$$\left(1 + \frac{1}{2\tau^2 \alpha_k^2 (\eta_k + 1)^2}\right) \|x^{k+1} - \bar{x}\|^2 \leq \left(\frac{1}{1 - 2\eta_k^2}\right) \|x^k - \bar{x}\|^2 \quad \forall k \geq \max\{k_0, \tilde{k}\},$$

lo cual implica que

$$\|x^{k+1} - \bar{x}\|^2 \leq \beta_k \|x^k - \bar{x}\|^2 \quad \forall k \geq \max\{k_0, \tilde{k}\}, \quad (6.16)$$

donde

$$\beta_k = \frac{1}{1 - 2\eta_k^2} \left(\frac{2\tau^2 (\eta_k + 1)^2}{2\tau^2 (\eta_k + 1)^2 + \frac{1}{\alpha_k^2}} \right). \quad (6.17)$$

Dado que $0 < \alpha_k \leq \bar{\alpha}$ para todo $k \in \mathbb{N}$, se tiene que

$$\beta_k \leq r_k, \quad (6.18)$$

donde

$$r_k = \frac{1}{1 - 2\eta_k^2} \left(\frac{2\tau^2(\eta_k + 1)^2}{2\tau^2(\eta_k + 1)^2 + \frac{1}{\bar{\alpha}^2}} \right).$$

Como $\eta_k \rightarrow 0$, en la expresión anterior tenemos que

$$r_k \rightarrow \frac{2\tau^2}{2\tau^2 + \frac{1}{\bar{\alpha}^2}} < 1.$$

Por lo tanto, existe un número positivo $k_1 \in \mathbb{N}$ con $k \geq k_1$, tal que

$$\beta_k \leq r_k < \frac{1}{2} \left(1 + \frac{2\tau^2}{2\tau^2 + \frac{1}{\bar{\alpha}^2}} \right) < 1 \quad \forall k \geq k_1.$$

Entonces, en (6.16) se tiene que

$$\|x^{k+1} - \bar{x}\|^2 \leq \beta_k \|x^k - \bar{x}\|^2 \quad \forall k \geq \max\{k_0, \tilde{k}, k_1\}.$$

Por lo tanto la sucesión $\{x^k\}$ converge linealmente a \bar{x} (ver Definición 2.1.11 (1)).

La convergencia superlineal de la sucesión generada por el algoritmo MPPEI2, se obtiene inmediatamente, dado que la hipótesis $\alpha_k \searrow 0$, y el hecho que $\eta_k \rightarrow 0$, en (6.17) implican que $\beta_k \rightarrow 0$, lo cual completa la prueba. \square

6.2. Experimentación numérica

En esta sección, con el objeto de mostrar la funcionalidad del MPPEI propuesto, presentamos dos ejemplos numéricos simples. Para ello usamos una computadora Intel Core(TM)2 Duo de 2.20 GHz y 3.00 GB de RAM, con sistema operativo Windows 10 de 64 bits, en el cuál implementamos nuestro código usando el software MATLAB 9.3.0 (R2017b).

Ejemplo 6.2.1. *Continuemos con el análisis del problema de minimización tratado en el Ejemplo 4.3.2, dado por:*

$$\min_{x \in \mathbb{R}} (e^x, (e^x - 3)^2 + 1). \quad (6.19)$$

Puesto que $F_1(x) = e^x$ y $F_2(x) = (e^x - 3)^2 + 1$ son funciones cuasi-conexas, entonces podemos decir que la función vectorial $F = (F_1, F_2)$ es \mathbb{R}_{\geq}^2 -Cuasi-conexa (ver Definición 2.2.11).

Este problema satisface también las hipótesis (H1), (H2) y (H3), y además como se probó en el Ejemplo 4.3.2, presenta a $X_E = (-\infty, \ln(3)] \approx (-\infty, 1.0986]$ como el conjunto de soluciones de Pareto del problema, también llamado conjunto

de soluciones eficientes.

Tomando a x^0 como punto inicial, y dado un $x^k \in \mathbb{R}^2$, el proceso iterativo del algoritmo propuesto (como se vio en la Observación 5.2.2), consiste en encontrar un punto crítico (mínimo local, máximo local o punto silla) del siguiente problema:

$$\min \left\{ \langle F(\cdot), z_k \rangle + \frac{\alpha_k}{2} \|\cdot - x^k\|^2 : x \in \Omega_k \right\}, \quad (6.20)$$

el cual también puede ser escrito como

$$\begin{cases} \min F(x) = (F_1(x), F_2(x))(z_1^k, z_2^k) + \frac{\alpha_k}{2} \|x - x^k\|^2 \\ \text{s.a. :} \\ (F_1(x), F_2(x)) = F(x) \leq F(x^k) = (F_1(x^k), F_2(x^k)). \end{cases}$$

Remplazando las componentes de F , tenemos

$$\begin{cases} \min F(x) = z_1^k(e^x) + z_2^k((e^x - 3)^2 + 1) + \frac{\alpha_k}{2} (x - x^k)^2 \\ \text{s.a. :} \\ e^x \leq e^{x^k} \\ (e^x - 3)^2 + 1 \leq (e^{x^k} - 3)^2 + 1. \end{cases}$$

Desarrollando las restricciones de desigualdad, y presentándolas en la forma ≤ 0 (lo cuál es un requerimiento de MATLAB), tenemos el siguiente programa de optimización:

$$\begin{cases} \min F(x) = z_1^k(e^x) + z_2^k((e^x - 3)^2 + 1) + \frac{\alpha_k}{2} (x - x^k)^2 \\ \text{s.a. :} \\ e^x - e^{x^k} \leq 0 \\ e^{2x} - e^{2x^k} + 6(e^{x^k} - e^x) \leq 0. \end{cases} \quad (6.21)$$

Asumiendo diferentes valores para el parámetro α_k , para $x^0 = 3$ (punto inicial para obtener la iteración cero), y para p^0 (punto inicial necesario para que MATLAB a través de la **función patternsearch** pueda resolver todos los subproblemas internos del problema (6.21)), vamos a mostrar que que la sucesión x^k generada converge hacia alguna de las soluciones de Pareto (también conocidas como conjunto de soluciones eficientes) del problema (6.19), las cuáles por cálculos previos, sabemos que se encuentran en el conjunto $X_E = (-\infty, \ln(3)]$ - ver Ejemplo 4.3.2.

Para ello vamos a considerar como fijos el parámetro $z_k = (z_1^k, z_2^k) = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$, el punto inicial $p^0 = 2$ (con el cuál se desarrollan todos los subproblemas), y $|x^{k+1} - x^k| < 0.0001$ como el criterio de parada para finalizar en algoritmo. A continuación, mostramos los resultados numéricos obtenidos al variar tanto el punto inicial x^0 , como el parámetro α_k .

1. Tomando a $x^0 = 3$ (como el punto inicial para obtener la iteración cero - punto que no pertenece al conjunto de soluciones eficientes $X_E = (-\infty, \ln(3)]$), y variando los parámetros α_k , tenemos:

1.1. Si $\alpha_k = 0.2$, para cada k , obtenemos:

k	$N[x^k]$	x^k	$ x^k - x^{k-1} $	$\sum F_i(x^k)z_i^k$	$F_1(x^k)$	$F_2(x^k)$
0	-	3	-	-	20.0855	297.9156
1	5	0.9595	2.0405	2.6603	2.6105	1.1517
2	3	0.9595	0.0000	2.6603	2.6105	1.1517

La tabla anterior muestra que con $k = 2$ iteraciones se resuelve el problema (observar que en este caso la sucesión $\{x^k\}$ converge al punto $x^2 \approx 0.9595 \in X_E = (-\infty, \ln(3)] \approx (-\infty, 1.0986]$). $N[x^k]$ denota el número de iteraciones internas de cada subproblema, con las cuáles se obtiene el punto x^k . Por ejemplo para obtener el punto $x^2 \approx 0.9595$, necesitamos $N[x^2] = 3$ iteraciones internas.

1.2. Si $\alpha_k = 0.6$, para cada k , obtenemos:

k	$N[x^k]$	x^k	$ x^k - x^{k-1} $	$\sum F_i(x^k)z_i^k$	$F_1(x^k)$	$F_2(x^k)$
0	-	3	-	-	20.0855	297.9156
1	5	1.0292	1.9708	2.7148	2.7988	1.0405
2	3	1.0292	0.0000	2.7148	2.7988	1.0405

La tabla anterior muestra que con $k = 2$ iteraciones se resuelve el problema.

1.3. Si $\alpha_k = 1$, para cada k , obtenemos:

k	$N[x^k]$	x^k	$ x^k - x^{k-1} $	$\sum F_i(x^k)z_i^k$	$F_1(x^k)$	$F_2(x^k)$
0	-	3	-	-	20.0855	297.9156
1	5	1.0845	0.9155	2.7999	2.9579	1.0018
2	3	1.0845	0.0000	2.7999	2.9579	1.0018

Equivalentemente a las tablas anteriores, la tabla anterior muestra que con $k = 2$ iteraciones, se obtiene el criterio de parada del problema, dado por $|x^2 - x^1| < 0.0001$.

En las tablas anteriores, podemos observar que al considerar en punto $x^0 = 3$ (el cuál no pertenece al conjunto de soluciones eficientes $X_E = (-\infty, \ln(3)]$ - calculado anteriormente en el Ejemplo 4.3.2), la sucesión $\{x^k\}$, haciendo uso de $k = 2$ iteraciones, converge hacia el punto $x^2 \approx 1 \in X_E$.

2. Tomando a $x^0 = -6$ (como el punto inicial para obtener la iteración cero - punto que es este caso pertenece al conjunto de soluciones eficientes $X_E = (-\infty, \ln(3)]$), y variando los parámetros α_k , tenemos:

2.1. Si $\alpha_k = 0.1$, para cada k , obtenemos:

k	$N[x^k]$	x^k	$ x^k - x^{k-1} $	$\sum F_i(x^k)z_i^k$	$F_1(x^k)$	$F_2(x^k)$
0	-	-6	-	-	0.0025	9.9855
1	3	-5.9999	0.0001	7.0623	0.0025	9.9855

La tabla anterior muestra que con $k = 1$ iteración se resuelve el problema. También nos muestra que para obtener el punto $x^1 = -5.9999 \in (-\infty, \ln(3)] = X_E$, necesitamos $N[x^1] = 3$ iteraciones internas.

Otro resultado semejante al mostrado en la tabla anterior, también se obtiene considerando $\alpha_k = 0.2, 0.3, \dots, 0.7$.

2.2. Si $\alpha_k = 0.8$, para cada k , obtenemos:

k	$N[x^k]$	x^k	$ x^k - x^{k-1} $	$\sum F_i(x^k)z_i^k$	$F_1(x^k)$	$F_2(x^k)$
0	-	-6	-	-	0.0025	9.9855
1	3	-6.0000	0.0000	7.0623	0.0025	9.9855

La tabla anterior nos muestra que necesitamos $k = 3$ iteraciones para resolver el problema. Considerando al parámetro $\alpha_k = 0.9$ y 1, también obtenemos un resultado semejante al mostrado en la tabla anterior.

En las dos tablas anteriores, podemos observar que al considerar el punto $x^0 = -6$ (el cuál pertenece al interior del conjunto de soluciones eficientes $X_E = (-\infty, \ln(3)]$ - calculado anteriormente en el Ejemplo 4.3.2), la sucesión $\{x^k\}$ converge (haciendo uso de una sola iteración) hacia el punto $x^1 \approx 6 \in X_E$.

3. Tomando a $x^0 = \ln(3) \approx 1.0986$ (como el punto inicial para obtener la iteración cero - punto que en este caso pertenece a la frontera del conjunto de soluciones eficientes $X_E = (-\infty, \ln(3)]$), y variando los parámetros α_k , tenemos:

3.1. Si $\alpha_k = 0.1$, para cada k , obtenemos:

k	$N[x^k]$	x^k	$ x^k - x^{k-1} $	$\sum F_i(x^k)z_i^k$	$F_1(x^k)$	$F_2(x^k)$
0	-	1.0986	-	-	3.0000	1.0000
1	4	1.0983	0.0003	2.8278	2.9991	1.0000
2	3	1.0982	0.0001	2.8275	2.9987	1.0000
3	3	1.0981	0.0001	2.8274	2.9985	1.0000

La tabla anterior muestra que con $k = 3$ iteraciones se resuelve el problema. Considerando que $\alpha_k = 0.2, 0.3, \dots, 1$, hemos obtenido un resultado semejante al mostrado en la tabla anterior. \square

En las tres situaciones anteriores (ítems 1, 2 y 3) en la que se asignaron diferentes valores para el punto inicial x^0 (necesario para obtener la iteración

cero), pertenecientes al interior, al complemento, y a la frontera del conjunto $X_E = (-\infty, \ln(3)]$, se pudo determinar la convergencia de la sucesión $\{x^k\}$ hacia algún punto perteneciente a dicho conjunto X_E .

Observación 6.2.1. Con el objeto de garantizar la convergencia del algoritmo, se ha modificado el Ejemplo 6.2.1, a la siguiente forma

$$\min_{\ln(2) \leq x \leq 10} (e^x, (e^x - 3)^2 + 1), \quad (6.22)$$

con lo cual se ha logrado reducir el intervalo que contiene al conjunto de soluciones de Pareto (calculado para el Ejemplo 6.2.1 como $X_E = (-\infty, \ln(3)]$), al siguiente conjunto $X_E = [\ln(2), \ln(3)] \approx [0.6931, 1.0986]$. Además, se ha logrado obtener un problema de minimización no irrestricto. Entonces, a pesar que el problema (6.22) no es irrestricto, probaremos que la sucesión generada por el algoritmo converge hacia alguna de las soluciones de Pareto $X_E = [\ln(2), \ln(3)]$.

A continuación, tomando un punto inicial x^0 (que no necesariamente pertenezca al intervalo $X_E = [\ln(2), \ln(3)]$), vamos a calcular que la sucesión $\{x^k\}$ converge hacia alguna de las soluciones de Pareto X_E de del problema (6.22). Para ello, vamos a considerar como fijos al parámetro $z_k = (z_1^k, z_2^k) = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$, y al punto inicial $x^0 = 3$ (necesario para obtener la iteración cero). Además, vamos a considerar a $|x^{k+1} - x^k| < 0.0001$ como el criterio de parada del algoritmo.

En las siguientes tablas, mostramos los resultados numéricos obtenidos al variar tanto el punto inicial p^0 , como el parámetro α_k .

1. Si $p^0 = 1$ y $\alpha_k = 0.2$, obtenemos:

k	$N[x^k]$	x^k	$ x^k - x^{k-1} $	$\sum F_i(x^k)z_i^k$	$F_1(x^k)$	$F_2(x^k)$
0	-	3	-	-	20.0855	292.9156
1	5	0.9596	2.0404	2.6603	2.6106	1.1517
2	2	0.9596	0.0000	2.6603	2.6106	1.1517

Asumiendo otros valores para el parámetro p^0 , como $p^0 = 2, 3, \dots, 7$, se obtienen resultados semejantes a los presentados en la tabla anterior, variando en este caso solamente el número de iteraciones internas $N[x^k]$.

2. Si $p^0 = 1$ y $\alpha_k = 0.5$, obtenemos:

k	$N[x^k]$	x^k	$ x^k - x^{k-1} $	$\sum F_i(x^k)z_i^k$	$F_1(x^k)$	$F_2(x^k)$
0	-	3	-	-	20.0855	292.9156
1	5	1.0134	1.9866	2.6976	2.7549	1.0601
2	2	1.0134	0.0000	2.6976	2.7549	1.0601

Esta tabla nos muestra que con $k = 2$ iteraciones se resuelve el problema.

3. Si $p^0 = 1$ y $\alpha_k = 1$, obtenemos:

k	$N[x^k]$	x^k	$ x^k - x^{k-1} $	$\sum F_i(x^k)z_i^k$	$F_1(x^k)$	$F_2(x^k)$
0	-	3	-	-	20.0855	292.9156
1	5	1.0845	1.9155	2.7999	2.9579	1.0018
2	2	1.0845	0.0000	2.7999	2.9579	1.0018

En las tres tablas anteriores, podemos observar que al considerar el punto factible $x^0 = 3$ (el cuál no pertenece al conjunto de soluciones eficientes $X_E = [\ln(2), \ln(3)] \approx [0.6931, 1.0986]$), la sucesión $\{x^k\}$ (haciendo uso de $k = 2$ iteraciones) converge hacia el punto $x^2 \approx 1 \in X_E$.

Ejemplo 6.2.2. Consideremos el siguiente problema de minimización:

$$\min_{x \in \mathbb{R}} (-e^{-x_1^2 - x_2^2} + 1, \sqrt{|x_1 - 1| + |x_2 - 2|}). \quad (6.23)$$

De forma análoga al ejemplo anterior, tenemos que la función vectorial $F = (F_1, F_2)$ es \mathbb{R}_{\geq}^2 -Cuasi-convexa, dado que las funciones $F_1(x) = -e^{-x_1^2 - x_2^2} + 1$ y $F_2(x) = \sqrt{|x_1 - 1| + |x_2 - 2|}$ son cuasi-convexas. Este problema también satisface las hipótesis **(H1)**, **(H2)** y **(H3)**. Además, presenta como soluciones de Pareto a los puntos $\bar{x} = (0, 0)$ y $\hat{x} = (1, 2)$.

Tomando a $x^0 = (x_1^0, x_2^0)$ como punto inicial, y dado un $x^k \in \mathbb{R}^2$, los pasos principales del algoritmo propuesto consisten en encontrar un punto crítico del siguiente problema:

$$\begin{cases} \min F(x_1, x_2) = z_1^k(-e^{-x_1^2 - x_2^2} + 1) + z_2^k(\sqrt{|x_1 - 1| + |x_2 - 2|}) + \\ \quad \frac{\alpha_k}{2} \|(x_1, x_2) - (x_1^k, x_2^k)\|^2 \\ \text{s.a :} \\ \quad -e^{-x_1^2 - x_2^2} + 1 \leq -e^{-(x_1^k)^2 - (x_2^k)^2} + 1 \\ \quad \sqrt{|x_1 - 1| + |x_2 - 2|} \leq \sqrt{|x_1^k - 1| + |x_2^k - 2|}, \end{cases}$$

el cuál en su forma estándar también puede ser escrito como:

$$\begin{cases} \min F(x_1, x_2) = z_1^k(-e^{-x_1^2 - x_2^2} + 1) + z_2^k(\sqrt{|x_1 - 1| + |x_2 - 2|}) + \\ \quad \frac{\alpha_k}{2} ((x_1 - x_1^k)^2 - (x_2 - x_2^k)^2) \\ \text{s.a :} \\ \quad x_1^2 + x_2^2 - (x_1^k)^2 - (x_2^k)^2 \leq 0 \\ \quad |x_1 - 1| + |x_2 - 2| - (|x_1^k - 1| + |x_2^k - 2|) \leq 0. \end{cases} \quad (6.24)$$

A continuación, asumiendo diferentes valores para α_k , probaremos que la sucesión x^k generada con la aplicación de la función patternsearch de MATLAB, converge hacia alguna de las soluciones de Pareto del problema.

Para ello vamos a considerar como fijos el parámetro $z_k = (z_1^k, z_2^k) = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$, el punto $x^0 = (-1, 3) = (x_1^0, x_2^0)$ con el cuál se obtiene la iteración 0, y el punto

inicial $p^0 = (2, 3) = (x_1, x_2)$ (necesario para que MATLAB a través de la **función patternsearch** pueda resolver todos los subproblemas internos de (6.24)). Además, consideraremos a $\|x^{k+1} - x^k\| < 0.0001$ como el criterio de parada para finalizar el algoritmo. Los resultados numéricos se muestran en las siguientes tablas:

1. Si $\alpha_k = 0.2$, para cada k , obtenemos:

k	$N[x^k]$	$x^k = (x_1^k, x_2^k)$	$\ x^k - x^{k-1}\ $	$\sum F_i(x^k)z_i^k$	$F_1(x^k)$	$F_2(x^k)$
0	-	$(-1, 3)$	-	-	0.9999	1.4142
1	5	$(1, 2)$	2.2361	0.7023	0.9933	0
2	3	$(1, 2)$	0	0.7023	0.9933	0

La tabla anterior muestra que con $k = 2$ iteración internas se resuelve el problema. $N[x^k]$ denota el número de iteraciones internas de cada subproblema, necesarias para obtener el punto x^k . Por ejemplo para obtener el punto $x^1 = (1, 2)$, necesitamos $N[x^1] = 5$ iteraciones internas.

2. Si $\alpha_k = 0.4$, para cada k , obtenemos:

k	$N[x^k]$	$x^k = (x_1^k, x_2^k)$	$\ x^k - x^{k-1}\ $	$\sum F_i(x^k)z_i^k$	$F_1(x^k)$	$F_2(x^k)$
0	-	$(-1, 3)$	-	-	0.9999	1.4142
1	2	$(-0.2713, 2.2259)$	1.063	1.5677	0.9934	1.2236
2	5	$(1, 2)$	1.2912	0.7023	0.9933	0
3	3	$(1, 2)$	0	0.7023	0.9933	0

La tabla anterior muestra que con $k = 3$ iteración se resuelve el problema. También nos muestra que para obtener, por ejemplo el el punto $x^1 = (x_1^k, x_2^k) = (-0.6016, 5.5977)$, necesitamos $N[x^1] = 1$ iteración interna.

3. Si $\alpha_k = 1$, para cada k , obtenemos:

k	$N[x^k]$	$x^k = (x_1^k, x_2^k)$	$\ x^k - x^{k-1}\ $	$\sum F_i(x^k)z_i^k$	$F_1(x^k)$	$F_2(x^k)$
0	-	$(-1, 3)$	-	-	0.9999	1.4142
1	2	$(-0.7785, 2.7776)$	0.3139	1.8374	0.9998	1.5988
2	5	$(-0.5309, 2.5258)$	0.3531	0.7203	0.9987	1.4341
3	5	$(-0.2339, 2.2078)$	0.4351	1.5510	0.9928	1.2007
4	3	$(0.1444, 2)$	0.4316	1.3485	0.9821	0.9250
5	3	$(0.4820, 1.9447)$	0.3421	1.2297	0.9819	0.7572
6	3	$(0.7857, 1.8418)$	0.3207	0.1258	0.9819	0.6103
7	4	$(1, 1.7337)$	0.2400	1.0591	0.9818	0.5160
8	3	$(1, 1.7337)$	0	1.0591	0.9818	0.5160

La tabla anterior nos muestra que con $k = 8$ iteración se obtiene el criterio de parada del problema.

En las tres tablas anteriores, podemos observar que la sucesión x^k converge hacia el punto $(1, 2)$, el cuál como se indicó al inicio de éste ejemplo, es una solución de Pareto del problema (6.23). \square

A continuación presentamos un ejemplo que no satisface la propiedad requerida de cuasi-convidad de F , y probamos que la sucesión $\{x^k\}$ generada por el algoritmo MPPEI, también converge hacia alguna solución de Pareto del mismo.

Ejemplo 6.2.3. Consideremos el siguiente problema de minimización:

$$\min_{x \in \mathbb{R}} (x^2, (x^2 - 2)^2). \quad (6.25)$$

Este problema satisface las hipótesis **(H1)** y **(H2)**, sin embargo no satisface la hipótesis **(H3)**, relacionada a la \mathbb{R}_{\leq}^p -cuasiconvexidad de $F = (x^2, (x^2 - 2)^2) = (F_1(x), F_2(x))$, dado que como podemos ver en la Figura 6.1, la segunda componente de la función F no es cuasi-convexa (observar la gráfica de $F_2(x)$ en el intervalo $[-\sqrt{2}, \sqrt{2}]$).

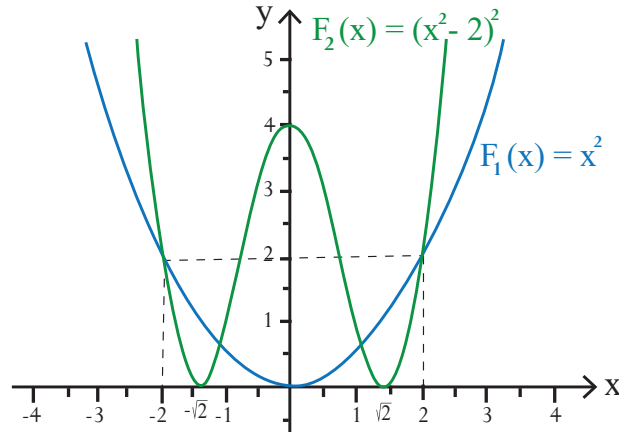


Figura 6.1: Gráfica de las funciones objetivo del Ejemplo 6.2.3.

Por cálculos análogos a los realizados en el Ejemplo 6.2.1 (ver lo indicado entre las relaciones (6.2) y (6.21)), se ha determinado que minimizar (6.25), a través del proceso iterativo dado por el algoritmo MPPEI, consiste en minimizar el siguiente problema:

$$\begin{cases} \min F(x) = z_1^k(x^2) + z_2^k(x^2 - 2)^2 + \frac{\alpha_k}{2}(x - x^k)^2 \\ \text{s.a. :} \\ x^2 - (x^k)^2 \leq 0 \\ (x^2 - 2)^2 - ((x^k)^2 - 2)^2 \leq 0. \end{cases}$$

Asumiendo diferentes valores para el parámetro α_k , y para el punto inicial x^0 (necesario para obtener la iteración cero), vamos a mostrar que la sucesión x^k generada con la aplicación de la función **patternsearch** de **MATLAB**, converge hacia alguna de las soluciones de Pareto del problema (6.25) contenidas en

el conjunto $X_E = [-\sqrt{2}, \sqrt{2}]$ (el cuál a sido determinado por cálculos análogos a los realizados en el Ejemplo 4.3.2).

Para ello vamos a considerar como fijos el parámetro $z_k = (z_1^k, z_2^k) = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ y el punto $x^0 = -1$ (con el cuál se obtiene la iteración 0). Además, vamos a considerar a $|x^{k+1} - x^k| < 0.0001$ como el criterio de parada para finalizar en algoritmo. Los resultados numéricos se muestran en las siguientes tablas:

1. Tomando a $x^0 = 5$ (como el punto inicial para obtener la iteración cero - punto que no pertenece al conjunto de soluciones eficientes $X_E = [-\sqrt{2}, \sqrt{2}]$), y variando los parámetros α_k , tenemos:

- 1.1. Si $\alpha_k = 0.3$, para cada k , obtenemos:

k	$N[x^k]$	x^k	$ x^k - x^{k-1} $	$\sum F_i(x^k)z_i^k$	$F_1(x^k)$	$F_2(x^k)$
0	-	5	-	-	25	529
1	5	1.3380	3.6620	1.2970	1.7903	0.0440
2	4	1.3380	0.0000	1.2970	1.7903	0.0440

La tabla anterior muestra que con $k = 2$ iteraciones se resuelve el problema. También nos muestra que para obtener el punto $x^2 \approx 1.3380 \in X_E$, necesitamos $N[x^2] = 3$ iteraciones internas.

- 1.2. Si $\alpha_k = 0.7$, para cada k , obtenemos:

k	$N[x^k]$	x^k	$ x^k - x^{k-1} $	$\sum F_i(x^k)z_i^k$	$F_1(x^k)$	$F_2(x^k)$
0	-	5	-	-	25	529
1	5	1.4510	3.5490	1.4966	2.1054	0.0111
2	4	1.3766	0.0744	1.3477	1.8949	0.0110
3	3	1.3766	0.0000	1.3477	1.8949	0.0110

La tabla anterior muestra que con $k = 3$ iteraciones se resuelve el problema.

- 1.3. Si $\alpha_k = 1$, para cada k , obtenemos:

k	$N[x^k]$	x^k	$ x^k - x^{k-1} $	$\sum F_i(x^k)z_i^k$	$F_1(x^k)$	$F_2(x^k)$
0	-	5	-	-	25	529
1	5	1.5197	3.4803	1.7009	2.3096	0.0959
2	4	1.3005	0.2193	1.2633	1.6912	0.0954
3	2	1.3005	0.0000	1.2633	1.6912	0.0954

La tabla anterior muestra que con $k = 3$ iteraciones se obtiene el criterio de parada del problema, dado por $|x^3 - x^1| < 0.0001$.

2. Tomando a $x^0 = -5$ (como el punto inicial para obtener la iteración cero - punto que tampoco pertenece al conjunto de soluciones eficientes $X_E = [-\sqrt{2}, \sqrt{2}]$), y variando los parámetros α_k , tenemos:

2.1. Si $\alpha_k = 0.2$, para cada k , obtenemos:

k	$N[x^k]$	x^k	$ x^k - x^{k-1} $	$\sum F_i(x^k)z_i^k$	$F_1(x^k)$	$F_2(x^k)$
0	-	-5	-	-	25	529
1	5	-1.3040	3.6960	1.2658	1.7004	0.0897
2	2	-1.3040	0.0000	1.2658	1.7004	0.0897

La tabla anterior muestra que con $k = 2$ iteraciones se resuelve el problema. Observar que la sucesión $\{x^k\}$ converge hacia el punto $x^2 \approx -1.3040 \in X_E = [-\sqrt{2}, \sqrt{2}] \approx [-1.4142, 1.4142]$.

2.2. Si $\alpha_k = 0.6$, para cada k , obtenemos:

k	$N[x^k]$	x^k	$ x^k - x^{k-1} $	$\sum F_i(x^k)z_i^k$	$F_1(x^k)$	$F_2(x^k)$
0	-	-5	-	-	25	529
1	5	-1.4255	3.5745	1.4376	2.0320	0.0010
2	4	-1.4029	0.0226	1.3924	1.9682	0.0010
3	3	-1.4029	0.0000	1.3924	1.9682	0.0010

La tabla anterior muestra que con $k = 3$ iteraciones se resuelve el problema.

2.3. Si $\alpha_k = 1$, para cada k , obtenemos:

k	$N[x^k]$	x^k	$ x^k - x^{k-1} $	$\sum F_i(x^k)z_i^k$	$F_1(x^k)$	$F_2(x^k)$
0	-	-5	-	-	25	529
1	5	1.5197	3.4803	1.7009	2.3096	0.0959
2	4	1.3005	0.2193	1.2633	1.6912	0.0954
3	2	1.3005	0.0000	1.2633	1.6912	0.0954

La tabla anterior muestra que con $k = 3$ se resuelve el problema. Observar que cuando el punto inicial x^0 (necesario para obtener la iteración cero), asume los valores $x^0 = \pm 5$, prácticamente las sucesiones respectivas convergen hacia los puntos $x^3 = \pm 1.3005$. \square

En las dos situaciones anteriores (ítems 1 y 2) en las cuáles se ha variado el punto inicial x^0 , asumiendo los valores $x^0 = \pm 5 \notin X_E = [-\sqrt{2}, \sqrt{2}]$, se ha podido determinar que la sucesión $\{x^k\}$ converge hacia un punto perteneciente al conjunto de soluciones eficientes X_E del problema 6.2.3.

Capítulo 7

METODOLOGÍA

7.1. Tipo y diseño de investigación

El presente trabajo es del tipo descriptivo - correlacional, pues en primera instancia describimos las propiedades de convergencia de los algoritmos de punto proximal para minimización mono-objetivo y multiobjetivo cuasi-convera, y en segunda instancia relacionamos dichos algoritmos, con su(s) version(es) inexacta(s).

7.2. Unidad de análisis

Para la realización de este trabajo, se estudió el artículo “A scalarization proximal point method for quasiconvex multiobjective minimization”, propuesto por Apolinário et al. (2016), y la tesis doctoral “Método do ponto proximal inexato para minimização quase-convera em variedades de Hadamard”, presentada por Baygorrea (2017) en la Universidad Federal de Rio de Janeiro.

7.3. Población de estudio

Por ser nuestro trabajo netamente abstracto, no existe población que estudiar, sin embargo nuestro estudio se encuentra íntimamente relacionada con el espacio Euclidiano \mathbb{R}^n , con los métodos de punto proximal, y con la optimización multiobjetivo cuasi-convera.

7.4. Técnicas de recolección de datos

El procedimiento de recolección de datos para la realización del presente trabajo (el cuál es netamente abstracto), básicamente consistió en la revisión bibliográfica (libros, artículos científicos, páginas web, tesis, etc.), en la asistencia a foros - eventos, y en la comunicación con investigadores.

7.5. Análisis e interpretación de la información

El avance del presente tesis fue gradual. Se inició este estudio revisando material referente a optimización mono-objetivo, el cual sirvió de base para adentrarse en el campo de la optimización multiobjetivo. Seguidamente se revisó teoría referente al análisis convexo, poniendo especial énfasis en las funciones cuasi-convexas y el cálculo subdiferencial. Paralelamente también se revisó algunos métodos y algoritmos de optimización, en especial el Método de Punto Proximal. Los resultados de dicha investigación fueron presentados semanalmente al asesor de la tesis quien orientó en el desarrollo de la misma.

Capítulo 8

RESULTADOS Y DISCUSIÓN

Presentación de resultados

El aporte del presente trabajo es de carácter teórico, pero con motivación a las aplicaciones. Los resultados de convergencia obtenidos por las dos versiones del algoritmo MPPEI, presentadas en el Capítulo 5, son novedosos en el espacio Euclidiano (incluso para el caso convexo el algoritmo es novedoso), y aún no han sido publicados por ninguna revista nacional o internacional. Entre los resultados más relevantes, destacamos los siguientes:

- 1. Se presenta un algoritmo MPPEI, y se prueba su convergencia global. Seguidamente, dotando de dos criterios de error al algoritmo, se obtiene dos versiones del mismo (el MPPEI1 y MPPEI2). Bajo hipótesis razonables, se prueba que la sucesión $\{x^k\}$ generada por cada una de éstas versiones converge hacia un punto crítico Pareto-Clarke del problema de optimización multiobjetivo (5.7).*
- 2. En el algoritmo MPPEI2, se a extendido el estudio de la tasa de convergencia de las funciones mono-objetivo ¹, a las multiobjetivo definidas en espacios Euclidianos, y bajo ciertas hipótesis naturales se ha probado que la sucesión x^k generada por dicho método, converge lineal y superlinealmente a un punto crítico Pareto-Clarke del problema (5.7), lo cual teóricamente es muy importante, pues nos da un mejor panorama de la eficiencia del método propuesto.*
- 3. Finalmente, con el objeto de mostrar la funcionalidad del algoritmo propuesto, y de las dos versiones del mismo, en el Capítulo 6 se presenta la implementación computacional de dos problemas de optimización, que satisfacen los requerimientos del algoritmo.*

Análisis, interpretación y discusión de resultados

- 1. Esta tesis puede ser considerada como un avance importante en la extensión del método de punto proximal para la resolución de problemas de optimi-*

¹Dicha tasa de convergencia fue introducida por Tang y Huang (2013), y posteriormente fue generalizada al caso cuasi-convexo por Baygorrea et al. (2016b).

zación no convexos, los cuáles según estudios recientes, modelan de una manera más aproximada ciertos problemas reales.

2. Los problemas que se modelan por funciones cuasi-convexas, por lo general son difíciles de resolver usando técnicas de optimización conocidas, es por ello que el aporte dado en esta tesis puede ser considerado como una contribución para la obtención de algoritmos de aproximación más eficientes.
3. Si en el algoritmo inicial MPPEI, en lugar del subdiferencial de Clarke $\partial^0(\cdot)$ hubiésemos considerado el subdiferencial en el límite $\partial^{Lim}(\cdot)$, habríamos obtenido la siguiente iteración

$$e^{k+1} \in \partial^{Lim} \left(\langle F(\cdot), z_k \rangle + \frac{\alpha_k}{2} \|\cdot - x^k\|^2 + \delta_{\Omega_k}(\cdot) \right) (x^{k+1}), \quad (8.1)$$

la cual desde el punto de vista teórico es más accesible para el estudio de la tasa de convergencia, dado que por ejemplo mejoraría la forma en que se ha definido el término g^k , calculado en la Observación 6.0.1.

Capítulo 9

CONCLUSIONES Y FUTURAS INVESTIGACIONES

9.1. Conclusiones

1. *En esta tesis, presentamos un Método de Punto Proximal Escalarizado Inexacto (MPPEI) para minimización de funciones multiobjetivo cuasi-convexas, definidas en espacios Euclidianos. Además, siguiendo un enfoque semejante al realizado por Baygorrea et al. (2016a), presentamos dos variantes inexactas del método: el MPPEI1 y MPPEI2, cada una de las cuáles es obtenida al agregar al MPPEI un criterio de error.*
2. *Bajo hipótesis razonables se prueba que la sucesión $\{x^k\}$ generada por cada una de las variantes del MPPEI, converge hacia algún punto crítico Pareto-Clarke del problema de optimización multiobjetivo (5.7). Además, bajo la asunción de ciertas hipótesis, se prueba que la sucesión generada por el algoritmo MPPEI2 converge lineal y superlinealmente hacia un punto crítico Pareto-Clarke.*
3. *El algoritmo MPPEI junto con sus dos versiones (MPPEI1 y MPPEI2) representan nuestro primer intento en la construcción de métodos de punto proximal eficientes para resolver problemas de minimización multiobjetivo cuasi-convexos.*
4. *Los métodos de aproximación inexactos, además de ajustarse mejor a la realidad, por lo general son más eficientes que los de búsqueda exacta, puesto que al considerar errores en su estructura, generan menos costo computacional, menor complejidad, y mejor tasa de convergencia.*
5. *Se espera que este trabajo de tesis sirva como referencia para futuros trabajos de investigación, pues hay otros temas por explorar, tanto en el aspecto teórico como computacional, dado su importancia en los distintos campos de la Ingeniería y Economía.*
6. *El Método de Punto Proximal (MPP) es un método iterativo de aproximación, que se caracteriza por que en cada iteración resuelve un modelo perturbado más simple, generando de esta manera una sucesión de puntos, los cuáles por lo general convergen a la solución de un determinado problema.*

9.2. Futuras investigaciones

En el transcurso de la realización de esta tesis, han surgido algunas cuestiones que no hemos resuelto y que esperamos desarrollar en futuros trabajos. Destacamos las siguientes:

- 1. Un futuro trabajo puede ser la implementación computacional del método estudiado para algunos problemas realistas, y su respectiva comparación con otros métodos.*
- 2. Estudiar si a través de la variación de los parámetros del algoritmo MP-PEI, es posible generar la frontera de Pareto, al igual que lo hacen otros métodos de optimización multiobjetivo, como el método de suma ponderada, el método de las ϵ - restricciones, etc.*
- 3. Otro trabajo importante también sería la extensión de dicho estudio, hacia el análisis de problemas de minimización multiobjetivo restrictos, los cuáles presentan un mayor rango de aplicación.*

REFERENCIAS BIBLIOGRÁFICAS

- Apolinário, H., Papa Quiroz, E., y Oliveira, P. (2016). *A scalarization proximal point method for quasiconvex multiobjective minimization*. Journal of Global Optimization, 64(1), 79–96.
- Attouch, H., y Bolte, J. (2009). *On the convergence of the proximal algorithm for nonsmooth functions involving analytic features*. Mathematical Programming, 116(1-2), 5–16.
- Attouch, H., Bolte, J., y Svaiter, B. (2013). *Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized gauss-seidel methods*. Mathematical Programming, 137(1-2), 91–129.
- Auslender, A., y Teboulle, M. (2006). *Interior gradient and proximal methods for convex and conic optimization*. SIAM Journal on Optimization, 16(3), 697–725.
- Ayres, F., y Mendelson, E. (1991). *Cálculo diferencial e integral*. Madrid: McGraw-Hill,.
- Baygorrea, N. (2017). *Método do ponto proximal inexato para minimização quase-convexa em variedades de Hadamard (Tesis Doctoral no publicada)*. Universidade Federal do Rio de Janeiro.
- Baygorrea, N., Papa Quiroz, E., y Maculan, N. (2016a). *Inexact proximal point methods for quasiconvex minimization on Hadamard manifolds*. Journal of the Operations Research Society of China, 4(4), 397–424.
- Baygorrea, N., Papa Quiroz, E., y Maculan, N. (2016b). *On the convergence rate of an inexact proximal point algorithm for quasiconvex minimization on Hadamard manifolds*. Journal of the Operations Research Society of China, 5(1), 1–11.
- Bazaraa, M., Sherali, H., y Shetty, C. (2013). *Nonlinear programming. theory and algorithms (3a. ed.)*. New Jersey: John Wiley & Sons.
- Benazic, R. (2000). *Topología en espacios Euclidianos (2a. ed.)*. Lima, Perú: Sociedad Matemática Peruana.
- Benazic, R. (2001). *Homeomorfismos en espacios Euclidianos*. Lima, Perú: Sociedad Matemática Peruana.

- Bento, G., Neto, J., y Soubeyran, A. (2014). *A proximal point-type method for multicriteria optimization*. Set-Valued and Variational Analysis, 22(3), 557–573.
- Berkovitz, L. (2003). Convexity and optimization in \mathbb{R}^n (Vol. 63). New York: John Wiley & Sons.
- Bolte, J., Daniilidis, A., Lewis, A., y Shiot, M. (2007). *Clarke subgradients of stratifiable functions*. SIAM Journal on Optimization, 18(2), 556–572.
- Bonnell, H., Iusem, A., y Svaiter, B. (2005). *Proximal methods in vector optimization*. SIAM Journal on Optimization, 15(4), 953–970.
- Ceng, L.-C., y Yao, J.-C. (2007). *Approximate proximal methods in vector optimization*. European Journal of Operational Research, 183(1), 1–19.
- Chen, J.-S., y Pan, S. (2008). *A proximal-like algorithm for a class of nonconvex programming*. Pacific Journal of Optimization, 4(2), 319–333.
- Clarke, F. (1990). Optimization and nonsmooth analysis (Vol. 5). Philadelphia: Siam.
- Clarke, F. (2013). Functional analysis, calculus of variations and optimal control (Vol. 264). New York: Springer Science & Business Media.
- Crouzeix, J., Ocaña, E., y Sosa, W. (2003). Análisis convexo. Lima, Perú: IMCA. (monografía N° 33)
- Cunha, F., da Cruz Neto, J., y Oliveira, P. (2010). *A proximal point algorithm with a ϕ -divergence for quasiconvex programming*. Optimization, 59(5), 777–792.
- Custódio, A., Madeira, J., Vaz, A., y Vicente, L. (2011). *Direct multisearch for multiobjective optimization*. SIAM Journal on Optimization, 21(3), 1109–1140.
- dos Santos Gromicho, J. (2013). Quasiconvex optimization and location theory (Vol. 9). Dordrecht, Holanda: Springer Science & Business Media.
- Ehrgott, M. (2005). Multicriteria optimization (2a. ed.). New York: Springer.
- Espinoza, E. (2004). sucesiones y series. Lima, Perú: Editorial Servicios Gráficos.
- Goudou, X., y Munier, J. (2009). *The gradient and heavy ball with friction dynamical systems: the quasiconvex case*. Mathematical Programming, 116(1-2), 173–191.
- Güler, O. (1992). *New proximal point proximal algorithms for convex minimization*. SIAM Journal on Optimization(2), 649-664.
- Kaplan, A., y Tichatschke, R. (1998). *Proximal point methods and nonconvex optimization*. Journal of global Optimization, 13(4), 389–406.

- Martinet, B. (1970). *Brève communication. régularisation d'inéquations variationnelles par approximations successives*. Revue française d'informatique et de recherche opérationnelle. Série rouge, 4(R3), 154–158.
- Mas-Colell, A., Whinston, M., y Green, J. (1995). *Microeconomic theory (Vol. 1)*. New York: Oxford university press.
- Mordukhovich, B. (2006). *Variational analysis and generalized differentiation I: Basic theory (Vol. 330)*. Berlin: Springer Science & Business Media.
- Navarro, F. (2013). *Algunas aplicaciones y extensión del método del subgradiente (Tesis de Master no publicada)*. Universidad Nacional Mayor de San Marcos.
- Pan, S., y Chen, J.-S. (2007). *Entropy-like proximal algorithms based on a second-order homogeneous distance function for quasi-convex programming*. Journal of Global Optimization, 39(4), 555–575.
- Papa Quiroz, E. (2017). *Optimización continua: teoría y ejercicios*. Lima, Perú: Editorial Académica Española.
- Papa Quiroz, E., y Oliveira, P. (2012). *Proximal point method for minimizing quasiconvex locally lipschitz functions on Hadamard manifolds*. Nonlinear Analysis: Theory, Methods & Applications, 75(15), 5924–5932.
- Papa Quiroz, E., Ramirez, L., y Oliveira, P. (2015). *An inexact proximal method for quasiconvex minimization*. European Journal of Operational Research, 246(3), 721–729.
- Pennanen, T. (2002). *Local convergence of the proximal point algorithm and multiplier methods without monotonicity*. Mathematics of Operations Research, 27(1), 170–191.
- Polyak, B. (1987). *Introduction to optimization*. Translations series in mathematics and engineering. New York.
- Rockafellar, R. (1970). *Convex analysis*. New Jersey: Princeton university press.
- Rockafellar, R. (1976). *Monotone operators and the proximal point algorithm*. SIAM journal on control and optimization(14), 877–898.
- Rockafellar, R., y Wets, R. (2009). *Variational analysis (Vol. 317)*. Berlin: Springer Science & Business Media.
- Sawaragi, Y., Nakayama, H., y Tanino, T. (1985). *Theory of multiobjective optimization*. mathematics in science and engineering (Vol. 176). Orlando, FL: Academic Press Inc.
- Souza, S., Oliveira, P., da Cruz Neto, J., y Soubeyran, A. (2010). *A proximal method with separable bregman distances for quasiconvex minimization over the nonnegative orthant*. European Journal of Operational Research, 201(2), 365–376.

- Tablada, C. (2011). Aproximación numérica equiespaciada de la variedad y el frente de Pareto para problemas de optimización multiobjetivo. *Recuperado de <https://rdu.unc.edu.ar/handle/11086/73>*.
- Takayama, A. (1995). *Mathematical economics (2a. ed.)*. New York: Cambridge University Press.
- Tang, G.-j., y Huang, N.-j. (2013). *An inexact proximal point algorithm for maximal monotone vector fields on Hadamard manifolds*. *Operations Research Letters*, 41(6), 586–591.
- Van Tiel, J. (1984). *Convex analysis: An introductory text*. Chichester, UK: John Wiley & Sons.
- Villacorta, K., y Oliveira, P. (2011). *An interior proximal method in vector optimization*. *European Journal of Operational Research*, 214(3), 485–492.